

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2002, 01:27:03 ; Search time 1893.2 seconds
(without alignments)
3614.503 Million cell updates/sec

Title: US-09-647-780A-3
Perfect score: 327
Sequence: 1 gggcagagatcacgcagcgg.....tcttcttcacactatgcc 327

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_btg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description

1	327	100.0	327	6	AX014703	Sequence 3 from Patent WO9553077
2	325.4	99.5	2636	6	AX139743	Sequence 13
3	325.4	99.5	2663	6	AX139745	Sequence 22
4	325.4	99.5	2676	6	AX033274	Sequence 29
5	325.4	99.5	2714	6	AX139741	Sequence 29
6	325.4	99.5	2784	9	AF336981	Sequence 29
7	323.8	99.0	2076	6	AX146976	Sequence 29
8	323.8	99.0	2262	6	AX146980	Sequence 29
9	323.8	99.0	2340	6	AX146978	Sequence 29
10	312.4	95.5	2232	6	AX319864	Sequence 29
11	237.4	72.6	2583	10	AF302075	Sequence 29
12	237.4	72.6	2652	10	AF302076	Sequence 29
13	237.4	72.6	2694	10	AF302077	Sequence 29
14	237.4	72.6	2925	6	AX033272	Sequence 29
15	237.4	72.6	2925	10	AF176569	Sequence 29
16	235.8	72.1	2601	10	AF157106	Sequence 29
17	235.8	72.1	2892	10	AF157105	Sequence 29
18	232.6	71.1	2765	6	AX014701	Sequence 29
19	170.8	52.2	3985	5	AF077612	Sequence 29
20	158	48.3	2557	9	HSCALLA	Sequence 29
21	158	48.3	3181	6	AX138810	Sequence 29
22	158	48.3	3181	6	AX138995	Sequence 29
23	158	48.3	3181	6	AX139414	Sequence 29
24	158	48.3	3181	6	I08056	Sequence 29
25	158	48.3	3181	9	HSMRNAEN	Sequence 29
26	158	48.3	5508	6	AX329921	Sequence 29
27	158	48.3	5508	6	AX333461	Sequence 29
28	158	48.3	5508	6	AX336695	Sequence 29
29	158	48.3	5508	9	HUMCALLA	Sequence 29
30	151	46.2	822	6	A30431	Sequence 29
31	151	46.2	2772	4	OCNEPR	Sequence 29
32	151	46.2	3243	6	I08057	Sequence 29
33	151	46.2	3243	10	RATENKA	Sequence 29
34	151	46.2	3369	6	E01557	Sequence 29
35	144.6	44.2	2253	10	MUSCD1ONEP	Sequence 29
36	141.8	43.4	154736	9	AL139246	Sequence 29
37	130	39.8	2340	10	S82653	Sequence 29
38	128.4	39.3	570	6	A44601	Sequence 29
39	128.4	39.3	2129	6	A44610	Sequence 29
40	128.4	39.3	2314	4	BTENDECE	Sequence 29
41	128.4	39.3	2314	6	A44617	Sequence 29
42	128.4	39.3	2610	4	S7377A	Sequence 29
43	127.4	39.0	2889	4	BTU27342	Sequence 29
44	127.4	39.0	2889	6	I74299	Sequence 29
45	126.2	38.6	1703	6	A44606	Sequence 29

ALIGNMENTS

RESULT	1	AX014703	327 bp	DNA	linear	PAT 07-SEP-2000
LOCUS	AX014703	Sequence 3 from Patent WO9553077				
DEFINITION	AX014703					
ACCESSION	AX014703.1	GI:10040977				
VERSION	human.					
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
TITLE	Schwartz,J.C., Gros,C., Ouimet,T., Rose,C., Bonhomme,M.C. and					
JOURNAL	Facchinetti,P.					
	Novel nep ii membrane metalloprotease and its use for screening					
	inhibitors useful in therapy					
	Patent: WO 9553077-A 3 21-OCT-1999;					
	INST NAT SANTE RECH MED (FR); SCHWARTZ JEAN CHARLES (FR); GROS					
	CLAUDE (FR); OUTMET TANJA (FR); ROSE CHRISTIANE (FR); BONHOMME					
	MARIE CHANTAL (FR); FACCHINETTI PATRICIA (FR)					
FEATURES	Location/Qualifiers					
source	1. 327					
	/organism="Homo sapiens"					

```

/db_xref="taxon:9606"
BASE COUNT 87 a 88 c 93 g 59 t
ORIGIN

Query Match 100.0%; Score 327; DB 6; Length 327;
Best Local Similarity 100.0%; Pred. No. 1.7e-68;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gggcacgagatcacgcacggctttgatgacaatggcggaaacttcgacaaagtggcaac 60
Db 1 GGGCAGCAGATCACGCACGGCTTTGATGACAATGGCGGAACCTTCGACAAAGATGGCAAC 60
Qy 61 atgatgattgtagagtaactctccaccagcacttcgggagcagctcagagtgcattg 120
Db 61 ATGATGATTGGTGGAGTAACCTCTCCACCAGCACCTTCGGGAGCAGCTCAGAGTGCATG 120
Qy 121 atctaccagtagcgcgaactactctctggacctggcagacgacgaaactgaaacggattc 180
Db 121 ATCTACCAGTAGCGCAACTACTCTCTGGACCTGGCAGACGAACAGCTGAACGGATTTC 180
Qy 181 aacaccttggggaaaaacattctgacaaacggaggggtgcggcaagcctataaagccctac 240
Db 181 AACACCTTGGGAAAAACATTCTGACAAACGAGGGGTGCGGCAAGCCTATAAGGCCTAC 240
Qy 241 ctcaagtggatggcagaggggtggaagaccagcagctgccggcctgagatctcacccat 300
Db 241 CTCAAGTGGATGGCAGAGGGGTGGAAGACCAGCAGCTGCCCGGCTGGATCTCACCCAT 300
Qy 301 gaggcagctcttcttcatcaactatgcc 327
Db 301 GAGCAGCTCTTCTTCATCACTATGCC 327

RESULT 2
AXI39743
LOCUS AXI39743 2636 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 3 from Patent EP1069188.
ACCESSION AXI39743
VERSION AXI39743.1 GI:14275325
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2636)
AUTHORS Jagerschmidt,A., Agnel,M. and Culouscou,J.M.
TITLE Three neprilysin-like membrane metalloproteinases
JOURNAL Patent: EP 1069188-A 3 17-JAN-2001;
SANOFI-SYNTHELABO (FR)
FEATURES
source
1..2636
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..7
8..2269
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:14275326"
/translation="MGKSEGPVGMVESAGRAGOKRPGFLEGLLLLLLLVTAALVALG
VLADRRGKIPAEQVSVCTPGCVIAAARILQNDPTPECDYFACGGLRRHV
IPETNSRYSIFDVLRLDEVLKAVLENSTAKPARVEKARTLRYSCMNQSVIKRGS
QPLLDILEVGGWVDMRWNETVGLWELEKALMNSQFNRRVLIDLFITNDQNS.
SRHYYIDQPTLQPSREYFNGSNRKYREAYLQFVSVATLLREDANLRFSLQV
EDMYVLEQLAKATVPOEERHDVIALYHRMGLLELQSQGLKGFDTLFIQTVLS
SVKIKLPDEEVVYGIPLYONLENIIDYTSARTIONLVRLDRTGSLSRFKDT
RVNTRKALFGTVEVRRECVCYVNSNENAGSLYREAFPGDSKSMVRELIDKVR
TVFVLELGLWDEESKKAQERKAMSIREQIGHDPYILEETNRRLDEEYENLFNSED
LYFNSLQNLKVGQSRSLRKLREKVDPLNLIIGAAVNFYSPNRNIVFPAGILOPP
FFSEQOALNFNGVIGHEITHGFDNGRNFKNMNMWNSFNSTFQREQSEC
MIYQYGNYSWLADEQNVNGTNGENIADNGVROAYKAYLKWAEQKQQLPGLD
LTHEQLFINAYQWCGSYREFFAIQSIKTDVHSPKLYRVLGSLQNLAAAFADTECAR
```

```

GTPMHPKRCRVW
2270..2635
BASE COUNT 601 a 759 c 813 g 462 t 1 others
ORIGIN

Query Match 99.5%; Score 325.4; DB 6; Length 2636;
Best Local Similarity 99.7%; Pred. No. 4.2e-68;
Matches 326; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gggcacgagatcacgcacggctttgatgacaatggcggaaacttcgacaaagtggcaac 60
Db 1763 GGGCAGCAGATCACGCACGGCTTTGACGACAATGGCGGAACCTTCGACAAAGATGGCAAC 1822
Qy 61 atgatgattgtagagtaactctccaccagcacttcgggagcagctcagagtgcattg 120
Db 1823 ATGATGATTGGTGGAGTAACCTCTCCACCAGCACCTTCGGGAGCAGCTCAGAGTGCATG 1882
Qy 121 atctaccagtagcgcgaactactctctggacctggcagacgacgaaactgaaacggattc 180
Db 1883 ATCTACCAGTAGCGCAACTACTCTCTGGACCTGGCAGACGAACAGCTGAACGGATTTC 1942
Qy 181 aacaccttggggaaaaacattctgacaaacggaggggtgcggcaagcctataaagccctac 240
Db 1943 AACACCTTGGGAAAAACATTCTGACAAACGAGGGGTGCGGCAAGCCTATAAGGCCTAC 2002
Qy 241 ctcaagtggatggcagaggggtggaagaccagcagctgccggcctgagatctcacccat 300
Db 2003 CTCAAGTGGATGGCAGAGGGGTGGAAGACCAGCAGCTGCCCGGCTGGATCTCACCCAT 2062
Qy 301 gaggcagctcttcttcatcaactatgcc 327
Db 2063 GAGCAGCTCTTCTTCATCACTATGCC 2089

RESULT 3
AXI39745
LOCUS AXI39745 2663 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 5 from Patent EP1069188.
ACCESSION AXI39745
VERSION AXI39745.1 GI:14275327
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2663)
AUTHORS Jagerschmidt,A., Agnel,M. and Culouscou,J.M.
TITLE Three neprilysin-like membrane metalloproteinases
JOURNAL Patent: EP 1069188-A 5 17-JAN-2001;
SANOFI-SYNTHELABO (FR)
FEATURES
source
1..2663
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..7
8..2296
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:14275328"
/translation="MGKSEGPVGMVESAGRAGOKRPGFLEGLLLLLLLVTAALVALG
VLADRRGKIPRLASRCFLQERTFVKRPRGIPAEQVSEVCTPGCVIAAARIL
QNDPTPECDYFACGGLRRHVIPETNSRYSIFDVLRLDEVLKAVLENSTAK
QNDPTPECDYFACGGLRRHVIPETNSRYSIFDVLRLDEVLKAVLENSTAK
PLHRAVEKARTLRYSCMNQSVIKRGSNRYREAYLQFVSVATLLREDANLRFSLQV
EDMYVLEQLAKATVPOEERHDVIALYHRMGLLELQSQGLKGFDTLFIQTVLS
SVKIKLPDEEVVYGIPLYONLENIIDYTSARTIONLVRLDRTGSLSRFKDT
RVNTRKALFGTVEVRRECVCYVNSNENAGSLYREAFPGDSKSMVRELIDKVR
TVFVLELGLWDEESKKAQERKAMSIREQIGHDPYILEETNRRLDEEYENLFNSED
LYFNSLQNLKVGQSRSLRKLREKVDPLNLIIGAAVNFYSPNRNIVFPAGILOPP
FFSEQOALNFNGVIGHEITHGFDNGRNFKNMNMWNSFNSTFQREQSEC
MIYQYGNYSWLADEQNVNGTNGENIADNGVROAYKAYLKWAEQKQQLPGLD
LTHEQLFINAYQWCGSYREFFAIQSIKTDVHSPKLYRVLGSLQNLAAAFADTECAR
```

KDQOLPGLDLTHEOLFFINVAQWCGSYRPEFAIQSIKTVDVHSPKRYLVLSLQNLAA
EADTFHCARGTTPMHPKRCRVW
2294..2563

BASE COUNT 609 a 769 c 816 g 468 t 1 others
ORIGIN

Query Match 99.5%; Score 325.4; DB 6; Length 2663;
Best Local Similarity 99.7%; Pred. No. 4.2e-68;
Matches 326; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggcacgagatcacgcacgcttttgatgacaattgcccgaacttcgacgaagatggcaac 60
DB 1790 GGGCAGGAGATCAGCAGCGCTTTGACGACAATGGCCGGAACCTTCGACAGAAATGGCAAC 1849

QY 61 atgatgattggtggagtaacttctccaccagcaacttcggggagcagtcagatgcatg 120
DB 1850 ATGATGATTGGTGGAGTAACCTTCTCCACCAGCACTTCGGGAGCAGTCAGAGTGCATG 1909

QY 121 atctaccagtagggcaactactctctgggaacttcggcagacgaacagcggattc 180
DB 1910 ATCTACCAGTACGGCACTACTCTCTGGAGCTTGGCAGACGAACAGACGTTGAACGGATTC 1969

QY 181 aacaccttggggaacattctgacacgaggggtgcccgaagcctataagccctac 240
DB 1970 AACACCTTGGGGAACATCTGTGACACGAGGGGTGCGGCAAGCTTAAAGCCCTAC 2029

QY 241 ctcaagtggatggcagaggggtggcaagaccagcagctgcccggcctggatctcaccat 300
DB 2030 CTCAAGTGGATGCGAGAGGTGGCAAGACCAGCAGCTGCCGGCCTGGATCTCACCCAT 2089

QY 301 gacgagctcttctcatcaactatgcc 327
DB 2090 GAGCAGCTCTTCTTCATCACTATGCC 2116

RESULT 4
AX033274 AX033274 2676 bp DNA linear PAT 21-SEP-2000
LOCUS Sequence 14 from Patent WO0047750.
DEFINITION AX033274
ACCESSION AX033274
VERSION AX033274.1 GI:10280089
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2676)
AUTHORS Boileau, G. and Desgroseillers, L.
TITLE New metalloproteases of the neprilysin family
JOURNAL Patent: WO 0047750-A 14 17-AUG-2000;
BOILEAU GUY (CA); DESGROSEILLERS LUC (CA); UNIVERSITE DE MONTREAL (CA)

FEATURES
source
1..2676
/organism="Homo sapiens"
/db_xref="taxon:9606"
7..2319
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09978.1"
/db_xref="GI:10280090"

/translation="MVESAGRAGKRPGLFEGGLLLLLVTAALVALGVLYADRRGK
OLPLASRLCFLQERTFVKRPGRGIPAEQVESEVCTTGGCVIAAARILQNMPTPEP
CDDFYQACGGLRRHVIPTNRSYISFDVLRDELEVLKAVLENSTAKRPAVEKAR
TLRSCNDSVIEKRSQPLDILEVGGMPVAMDRWNETVGLWELEWELRLALNWSOF
NRVLDFLEINDDQNSRHIIYIDPTGMPSPREYFNGGSRKREAYLQPMVSV
TLREDANPLRDSCLVQDMVQVLETLQAKATVQERHDIYALYHRMGLBELQSQ
FLGKGNWTLFIOTLVSSVKIKLLPDEEVVYGIPLYLQNLNIDITYSARTIONLVW
RLVLDIGLSORFKDTRVNYRKALFCTMVVEVRRECVCYVNSNMENAVGSLYVREA
FPDGSKSMVRLDKVTVFVETDELGMWDESKKKAQKAMSIREQIGHDYLILE
MNRLEDEYSNLNIDITYSARTIONLVWRLVLDIGLSORFKDTRVNYRKALFCTMVVEVRRECVCYVNSNMENAVGSLYVREA
SPNRNIVFPAGILOPPFTSKQEPQALNFGGIGWIGHEITHGFDNDGRNFGNMM

DWMSNESTOHFREQSECMYIOYGNYSMDLADENNVNGNTLGENIADNGVROAYKAY
LKNWEGKQDQPLGLDLTHEOLFFINVAQWCGSYRPEFAIQSIKTVDVHSPKRYLV
GSLQNLAAAFADTFHCARGTTPMHPKRCRVW
608 a 771 c 823 g 474 t

BASE COUNT
ORIGIN

Query Match 99.5%; Score 325.4; DB 6; Length 2676;
Best Local Similarity 99.7%; Pred. No. 4.2e-68;
Matches 326; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggcacgagatcacgcacgcttttgatgacaattgcccgaacttcgacgaagatggcaac 60
DB 1813 GGGCAGGAGATCAGCAGCGCTTTGACGACAATGGCCGGAACCTTCGACAGAAATGGCAAC 1872

QY 61 atgatgattggtggagtaacttctccaccagcaacttcggggagcagtcagatgcatg 120
DB 1873 ATGATGATTGGTGGAGTAACCTTCTCCACCAGCACTTCGGGAGCAGTCAGAGTGCATG 1932

QY 121 atctaccagtagggcaactactctctgggaacttcggcagacgaacagcggattc 180
DB 1933 ATCTACCAGTACGGCACTACTCTCTGGAGCTTGGCAGACGAACAGACGTTGAACGGATTC 1992

QY 181 aacaccttggggaacattctgacacgaggggtgcccgaagcctataagccctac 240
DB 1993 AACACCTTGGGGAACATCTGTGACACGAGGGGTGCGGCAAGCTTAAAGCCCTAC 2052

QY 241 ctcaagtggatggcagaggggtggcaagaccagcagctgcccggcctggatctcaccat 300
DB 2053 CTCAAGTGGATGCGAGAGGTGGCAAGACCAGCAGCTGCCGGCCTGGATCTCACCCAT 2112

QY 301 gacgagctcttctcatcaactatgcc 327
DB 2113 GAGCAGCTCTTCTTCATCACTATGCC 2139

RESULT 5
AX139741 AX139741 2714 bp DNA linear PAT 30-MAY-2001
LOCUS Sequence 1 from Patent EP1069188.
DEFINITION AX139741
ACCESSION AX139741
VERSION AX139741.1 GI:14275323
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2714)
AUTHORS Jägerschmidt, A., Agnel, M. and Culouscou, J. M.
TITLE Three neprilysin-like membrane metalloproteinases
JOURNAL Patent: EP 1069188-A 1 17-JAN-2001;
SANOFI-SYNTHELABO (FR)

FEATURES
source
1..2714
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..7
8..2347
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC39974.1"
/db_xref="GI:14275324"

/translation="MKGSEGVGMVESAGRAGKRPGLFEGGLLLLLVTAALVALG
VLVADRQKOLPLASRLCFLQERTFVKRPGRGIPAEQVESEVCTTGGCVIAAARIL
QNMPTPEPDDDFQACGGLRRHVIPTNRSYISFDVLRDELEVLKAVLENSTAK
DRAVEKARTLIRSCNDSVIEKRSQPLDILEVGGMPVAMDRWNETVGLWELEWELRLALNWSOF
NRVLDFLEINDDQNSRHIIYIDPTGMPSPREYFNGGSRKREAYLQPMVSV
YLAQPMVSVATLREDANPLRDSCLVQDMVQVLETLQAKATVQERHDIYALYHRMGLBELQSQ
FLGKGNWTLFIOTLVSSVKIKLLPDEEVVYGIPLYLQNLNIDITYSARTIONLVW
RTIONLVWRLVLDIGLSORFKDTRVNYRKALFCTMVVEVRRECVCYVNSNMENAVGSLYVREA
VGSILYVREAPGDSKSMVRLDKVTVFVETDELGMWDESKKKAQKAMSIREQIGHDYLILE
GHDPYILLEETNRLEDEYSNLNIDITYSARTIONLVWRLVLDIGLSORFKDTRVNYRKALFCTMVVEVRRECVCYVNSNMENAVGSLYVREA
GAAVYNVAFYSPNRNIVFPAGILOPPFTSKQEPQALNFGGIGWIGHEITHGFDNDGR


```
/db_xref="GI:14346248"
/translation="CTTGGCCTTTCCTTCATCACTAGGCC 1896
VLYADRRGIPEAQEVSEVCTTGGCCTTTCCTTCATCACTAGGCC 1896
IPETSRYSIFDLVLRDELEVLKAVLENSTADKPAVERKARTLYRSCMQSVLEKRG
QPLLDILEVVGWPMVDRNMTVGLWELERQALMNSQFNRRVLIDLFIMNDQNS
SHIYIDOPTLGMPSREYFNGGSKNRKREAYLQFMVSVATLLREDANLPRDSCVQ
EDMOVLLETOLAKATVPOERHDVIALYHRMGLLELOSOFLKGFNTLFTQVLVS
SVKILLPDEEVVVGIPYLNLENIDITYSARTIONYLVRVLDRIGSLRSOKRDT
RVNPKALFGTMVEEVRRECVCYVNSNMENAVSILYREAFPGDSKSMVRELIDKVR
TVFETDLGMDDESKKAKKAKMSIREQIGHDPYIILEMNRRLDEEYSLNEDYFV
ETDELGMDESKKAKKAKMSIREQIGHDPYIILEMNRRLDEEYSLNEDYFV
BQALNKGAGRSURKREKVDNLIIGAAVNAFYSPNRNQIYFPAGILQPPFSK
NSPOLNFGAGTGMVIGHEITHGFDNGRDKNGNMNDWNSFSTQHFREQSECMIY
YGNLSDWLDONVNGFNTLGENIADNGVRQAYKAYLKMAEGGKDOOLPCLDITHE
QLEFNTYAOVWCGSVYRPEFAIQSIKTDVHSPKLYRVLSLQNLAAFADEFFHCARTPM
HPKRECRVW"
BASE COUNT 493 a 578 c 636 g 369 t
ORIGIN
```

```
Query Match 99.0%; Score 323.8; DB 6; Length 2076;
Best Local Similarity 99.4%; Pred. No. 1e-67;
Matches 325; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gggcagagatcacgcagcgtttgatgacaatggcgggaacttcgacaagaatggcaac 60
Db 1570 GGGCAGGATCACGCAGCGCTTTCGACGACATGGCGGAATTCGACGAATGGCAAC 1629

Qy 61 atgatggattgaggaatctctccaccagcacttcgggagcagtcagatgcatg 120
Db 1630 ATGATGGATTGAGGAATCTCTCCACCAGCAGCTTCGGGAGCAGTCAGATGCGATG 1689

Qy 121 atctaccagtcgcaactctctggagcctggcagacagacagacagcagcagc 180
Db 1690 ATCTACAGTCGCGCAACTCTCTGGGACCTTGGCAGACGACAGAACGTCGAACGGATTC 1749

Qy 181 aacacccctggggaaacattctgacaacagaggggtgcggcaagcctataagggctac 240
Db 1750 AACACCCCTGGGGAAACATTCTGACACGAGGAGGGGTGGCGCAACCTATAGGCCTAC 1809

Qy 241 ctcaagtggatggcagaggggtggcaggaaccagcagctccggcgctgagatctacccat 300
Db 1810 CTCAAGTGGATGGCAGAGGGTGGCAGGACGACGAGCTCCCGCGCTGGATCTCACCCAT 1869

Qy 301 gagcagctctcttcataactatgcc 327
Db 1870 GAGCAGCTCTTCTTCATCACTAGGCC 1896
```

```
RESULT 8
AXI46980 AXI46980 2262 bp DNA linear PAT 08-JUN-2001
LOCUS Sequence 5 from Patent WO0136610.
DEFINITION ACCESSION AXI46980
VERSION AXI46980.1 GI:14346251
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Deleersnijder W., Wiegers R. and Weske M.
TITLE Human enzymes of the metalloprotease family
JOURNAL Patent: WO 0136610-A 5 25-MAY-2001;
Solvay Pharmaceuticals B.V. (NL)
FEATURES
source
1..2262
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS
1..2262
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC41161.1"
/db_xref="GI:14346251"
/translation="MKGSEGPVGMVESAGRAGQKRPGLFEGGLLLLLLVTAALVALG
VLYADRRGIPEAQEVSEVCTTGGCCTTTCCTTCATCACTAGGCC 1896
VLYADRRGIPEAQEVSEVCTTGGCCTTTCCTTCATCACTAGGCC 1896
IPETSRYSIFDLVLRDELEVLKAVLENSTADKPAVERKARTLYRSCMQSVLEKRG
QPLLDILEVVGWPMVDRNMTVGLWELERQALMNSQFNRRVLIDLFIMNDQNS
SHIYIDOPTLGMPSREYFNGGSKNRKREAYLQFMVSVATLLREDANLPRDSCVQ
EDMOVLLETOLAKATVPOERHDVIALYHRMGLLELOSOFLKGFNTLFTQVLVS
SVKILLPDEEVVVGIPYLNLENIDITYSARTIONYLVRVLDRIGSLRSOKRDT
RVNPKALFGTMVEEVRRECVCYVNSNMENAVSILYREAFPGDSKSMVRELIDKVR
TVFETDLGMDDESKKAKKAKMSIREQIGHDPYIILEMNRRLDEEYSLNEDYFV
ETDELGMDESKKAKKAKMSIREQIGHDPYIILEMNRRLDEEYSLNEDYFV
BQALNKGAGRSURKREKVDNLIIGAAVNAFYSPNRNQIYFPAGILQPPFSK
NSPOLNFGAGTGMVIGHEITHGFDNGRDKNGNMNDWNSFSTQHFREQSECMIY
YGNLSDWLDONVNGFNTLGENIADNGVRQAYKAYLKMAEGGKDOOLPCLDITHE
QLEFNTYAOVWCGSVYRPEFAIQSIKTDVHSPKLYRVLSLQNLAAFADEFFHCARTPM
HPKRECRVW"
```

```
VLYADRRGIPEAQEVSEVCTTGGCCTTTCCTTCATCACTAGGCC 1896
IPETSRYSIFDLVLRDELEVLKAVLENSTADKPAVERKARTLYRSCMQSVLEKRG
QPLLDILEVVGWPMVDRNMTVGLWELERQALMNSQFNRRVLIDLFIMNDQNS
SHIYIDOPTLGMPSREYFNGGSKNRKREAYLQFMVSVATLLREDANLPRDSCVQ
EDMOVLLETOLAKATVPOERHDVIALYHRMGLLELOSOFLKGFNTLFTQVLVS
SVKILLPDEEVVVGIPYLNLENIDITYSARTIONYLVRVLDRIGSLRSOKRDT
RVNPKALFGTMVEEVRRECVCYVNSNMENAVSILYREAFPGDSKSMVRELIDKVR
TVFETDLGMDDESKKAKKAKMSIREQIGHDPYIILEMNRRLDEEYSLNEDYFV
ETDELGMDESKKAKKAKMSIREQIGHDPYIILEMNRRLDEEYSLNEDYFV
BQALNKGAGRSURKREKVDNLIIGAAVNAFYSPNRNQIYFPAGILQPPFSK
NSPOLNFGAGTGMVIGHEITHGFDNGRDKNGNMNDWNSFSTQHFREQSECMIY
YGNLSDWLDONVNGFNTLGENIADNGVRQAYKAYLKMAEGGKDOOLPCLDITHE
QLEFNTYAOVWCGSVYRPEFAIQSIKTDVHSPKLYRVLSLQNLAAFADEFFHCARTPM
HPKRECRVW"
BASE COUNT 520 a 628 c 716 g 398 t
ORIGIN
```

```
Query Match 99.0%; Score 323.8; DB 6; Length 2262;
Best Local Similarity 99.4%; Pred. No. 1e-67;
Matches 325; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gggcagagatcacgcagcgtttgatgacaatggcgggaacttcgacaagaatggcaac 60
Db 1756 GGGCAGGATCACGCAGCGCTTTCGACGACATGGCGGAATTCGACGAATGGCAAC 1815

Qy 61 atgatggattgaggaatctctccaccagcacttcgggagcagtcagatgcatg 120
Db 1816 ATGATGGATTGAGGAATCTCTCCACCAGCAGCTTCGGGAGCAGTCAGATGCGATG 1875

Qy 121 atctaccagtcgcaactctctggagcctggcagacagacagacagcagcagc 180
Db 1876 ATCTACAGTCGCGCAACTCTCTGGGACCTTGGCAGACGACAGAACGTCGAACGGATTC 1935

Qy 181 aacacccctggggaaacattctgacaacagaggggtgcggcaagcctataagggctac 240
Db 1936 AACACCCCTGGGGAAACATTCTGACACGAGGAGGGGTGGCGCAACCTATAGGCCTAC 1995

Qy 241 ctcaagtggatggcagaggggtggcaggaaccagcagctccggcgctgagatctacccat 300
Db 1996 CTCAAGTGGATGGCAGAGGGTGGCAGGACGACGAGCTCCCGCGCTGGATCTCACCCAT 2055

Qy 301 gagcagctctcttcataactatgcc 327
Db 2056 GAGCAGCTCTTCTTCATCACTAGGCC 2082
```

```
RESULT 9
AXI46978 AXI46978 2340 bp DNA linear PAT 08-JUN-2001
LOCUS Sequence 3 from Patent WO0136610.
DEFINITION ACCESSION AXI46978
VERSION AXI46978.1 GI:14346249
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Deleersnijder W., Wiegers R. and Weske M.
TITLE Human enzymes of the metalloprotease family
JOURNAL Patent: WO 0136610-A 3 25-MAY-2001;
Solvay Pharmaceuticals B.V. (NL)
FEATURES
Location/Qualifiers
1..2340
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS
<1..2340
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC41160.1"
/db_xref="GI:14346250"
/translation="MKGSEGPVGMVESAGRAGQKRPGLFEGGLLLLLLVTAALVALG
VLYADRRGIPEAQEVSEVCTTGGCCTTTCCTTCATCACTAGGCC 2082
VLYADRRGIPEAQEVSEVCTTGGCCTTTCCTTCATCACTAGGCC 2082
IPETSRYSIFDLVLRDELEVLKAVLENSTADKPAVERKARTLYRSCMQSVLEKRG
QPLLDILEVVGWPMVDRNMTVGLWELERQALMNSQFNRRVLIDLFIMNDQNS
SHIYIDOPTLGMPSREYFNGGSKNRKREAYLQFMVSVATLLREDANLPRDSCVQ
EDMOVLLETOLAKATVPOERHDVIALYHRMGLLELOSOFLKGFNTLFTQVLVS
SVKILLPDEEVVVGIPYLNLENIDITYSARTIONYLVRVLDRIGSLRSOKRDT
RVNPKALFGTMVEEVRRECVCYVNSNMENAVSILYREAFPGDSKSMVRELIDKVR
TVFETDLGMDDESKKAKKAKMSIREQIGHDPYIILEMNRRLDEEYSLNEDYFV
ETDELGMDESKKAKKAKMSIREQIGHDPYIILEMNRRLDEEYSLNEDYFV
BQALNKGAGRSURKREKVDNLIIGAAVNAFYSPNRNQIYFPAGILQPPFSK
NSPOLNFGAGTGMVIGHEITHGFDNGRDKNGNMNDWNSFSTQHFREQSECMIY
YGNLSDWLDONVNGFNTLGENIADNGVRQAYKAYLKMAEGGKDOOLPCLDITHE
QLEFNTYAOVWCGSVYRPEFAIQSIKTDVHSPKLYRVLSLQNLAAFADEFFHCARTPM
HPKRECRVW"
```



```

source      1. .2694
/organism="Mus musculus"
/db_xref="taxon:10090"
25. .2364
/notes="endopeptidase"
/codon_start=1
/protein_id="AAG18448.1"
/product="neprilysin-like peptidase gamma"
/db_xref="GI:10505364"
/translation="MVERAGWCKKSPGFVEYGLMVLILLGLAIVTLGVFYSIGKOL
SLKSDICTPSCVIAARILNMDQSRNPFYQACGWLRIHVIPETNRSYVF
DILRDEVLKGLVDSLSOHPAPVEKAKTLYRSCMNOSVIEKRDSEPLLVKMGV
GWPMKWNTEKGLWELEROLAVLNQFNRRVLIDLFTWDDQSSRHVYIDOPT
LGMPSREYFOEDNNKVRKAYLEFMTSVATMLRKDONLSKESAMVREMAEVLLET
HLANATVPQEKRDHVTALYHRMDLMELOERGLKDRVSLCSPCCPGTHSDQAGLELG
NPPASRVGLGKWNFTILQVLSVEFLPDEEVYVIGPILENLEDDIIDSYSIA
RTMNYLWRLVLDRLGSKAEKRDYRKALYGTVEEVREKCVSNMESIA
VSLYIKRAFSKDSKTSRELIKRSVFDNLNMDDESKKAEKAMNIRQI
GYPDYILEDNKHLDEYSLSLTYEDLYFENGONLKNNAORSLKLRKVDONLWII
GAUVNAFYSPNNOIVFPAGILOPFFSKDQPSLNFGGIGMVGHEITHGFDNGR
NFDKNGMLDWSNFSARHFOQQSCMIYQGNFSELADNQNNGFSTLGENIADNG
GWQAYKAYLRWLADGGKQDRLPLNLTYAOLFFINYAQMCGSYRPEFAVOSIKTDV
HSPKRYVLGSLQNLPGFSEAFHCPGSPMHPMKRCRIW"
BASE COUNT      686 a 700 c 766 g 542 t
ORIGIN

Query Match      72.6%; Score 237.4; DB 10; Length 2694;
Best Local Similarity 82.9%; Pred. No. 5.8e-47;
Matches 271; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1 gggcacgagatcacgcagcggctttgatgacaatggccgggaacttcgacaagaatggcaac 50
DB 1858 GGGCACGAGATCACACACGGCTTTGATGATAATGCTGCTAACTTTGACAAGACGGCAAC 1317

QY 61 atgatgattggtgagtaactctccaccagcactccggggagcagtcagatgcatg 120
DB 1918 ATGCTGGACTGGTGGAGTAACCTTCGCGCCGGCAGCTTCCAAACAGCAGTCGCAATGATG 1977

QY 121 atctaccagtcagcgaactactcctcctgggaactggcagacgaacagcgtgaacgattc 180
DB 1978 ATCTATCAGTACGCAACTTCTCTGGGAACATAGCAGACCAACAGAGATGTGAACGGATTC 2037

QY 181 aacaccttggggaaacattgtgacaacggagggtggcggaagcctataaagccctac 240
DB 2038 AGTACCTCGGGGAGACATTGCCGACAAACGAGGTGTGCGACAGGCATACAAAGCTTAC 2097

QY 241 ctcaagtggatggcagaggggtggcaagaccagcagctgcccggcctggtatctcacccat 300
DB 2098 CTACGGTGGCTGGCTGATGGCGGCAAGATCAGCGACTGCCGGGACTGAACCTGACCTAT 2157

QY 301 gaggcagctcttcttcatacaactatgcc 327
DB 2158 GCCCAGCTTTTCTTCATCACTATGCC 2184

RESULT 14
AX033272
LOCUS      2925 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 12 from Patent WO0047750.
ACCESSION AX033272
VERSION    AX033272.1 GI:10280087
KEYWORDS   house mouse.
SOURCE     Mus musculus
ORGANISM   Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2925)
AUTHORS   Boileau,G. and Desgroseillers,L.
TITLE     New metalloproteases of the neprilysin family
JOURNAL   Patent: WO 0047750-A 12 17-AG-2000;
BOILEAU GUY (CA); DESGROSEILLERS LUC (CA); UNIVERSITE DE MONTREAL
(CA)

```

```

Location/Qualifiers
1. .2925
/organism="Mus musculus"
/db_xref="taxon:10090"
332. .2629
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC09977.1"
/db_xref="GI:10280088"
/translation="MVERAGWCKKSPGFVEYGLMVLILLGLAIVTLGVFYSIGKOL
PLTSLHSDERTVVKRALRDSSLSKSDICTPSCVIAARILNMDQSRNPFY
CMQAGWLRIHVIPETNRSYVFEDILRDEVLKGLVDSLSOHPAPVEKAKTLYRS
CMQSVTEKRDSEPLLVKMGVPMKWNTEKGLWELEROLAVLNQFNRRVL
IDLFTWDDQSSRHVYIDOPTLGMPSREYFOEDNNKVRKAYLEFMTSVATMLRK
DONLSKESAMVREMAEVLLETHLANATVPQEKRDHVTALYHRMDLMELOERGLK
FNWLTQNLVSSVEFLPDEEVYVIGPILENLEDDIIDSYSIAKTNLWRLVLD
RIGSLORFEKRDYRKALYGTVEEVREKCVSNMESIAVSLYIKRAFSKDS
KTSVRELIERKRSVFDNLNMDDESKKAEKAMNIRQIYQGNFSELADNQNNGF
DEYSSLTIFYEDLYFENGONLKNNAORSLKLRKVDONLWIIIGAAVNAFYSPNRN
QIVFPAGILOPFFSKDQPSLNFGGIGMVGHEITHGFDNGRNFDRKNGMLDWSN
FSARHFOQQSCMIYQGNFSELADNQNNGFSTLGENIADNGGVQAYKAYLRWL
LPGKQDRLPLNLTYAOLFFINYAQMCGSYRPEFAVOSIKTDVHSPKRYVLGSLQ
LPQFSEAFHCPGSPMHPMKRCRIW"
BASE COUNT      710 a 797 c 836 g 582 t
ORIGIN

Query Match      72.6%; Score 237.4; DB 6; Length 2925;
Best Local Similarity 82.9%; Pred. No. 5.8e-47;
Matches 271; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1 gggcacgagatcacgcagcggctttgatgacaatggccgggaacttcgacaagaatggcaac 60
DB 2123 GGGCACGAGATCACACACGGCTTTGATGATAATGCTGCTAACTTTGACAAGACGGCAAC 2182

QY 61 atgatgattggtgagtaactctccaccagcactccggggagcagtcagatgcatg 120
DB 2183 ATGCTGGACTGGTGGAGTAACCTTCGCGCCGGCAGCTTCCAAACAGCAGTCGCAATGATG 2242

QY 121 atctaccagtcagcgaactactcctcctgggaactggcagacgaacagcgtgaacgattc 180
DB 2243 ATCTATCAGTACGCAACTTCTCTGGGAACATAGCAGACCAACAGAGATGTGAACGGATTC 2302

QY 181 aacaccttggggaaacattgtgacaacggagggtggcggaagcctataaagccctac 240
DB 2303 AGTACCTCGGGGAGACATTGCCGACAAACGAGGTGTGCGACAGGCATACAAAGCTTAC 2362

QY 241 ctcaagtggatggcagaggggtggcaagaccagcagctgcccggcctggtatctcacccat 300
DB 2363 CTACGGTGGCTGGCTGATGGCGGCAAGATCAGCGACTGCCGGGACTGAACCTGACCTAT 2422

QY 301 gaggcagctcttcttcatacaactatgcc 327
DB 2423 GCCCAGCTTTTCTTCATCACTATGCC 2449

RESULT 15
AF176569
LOCUS      2925 bp mRNA linear ROD 11-MAY-2000
DEFINITION Mus musculus neprilysin-like metalloproteinase 1 (N11) mRNA,
complete cds.
ACCESSION AF176569
VERSION    AF176569.1 GI:7769082
KEYWORDS   house mouse.
SOURCE     Mus musculus
ORGANISM   Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2925)
AUTHORS   Ghaddar,G., Ruchon,A.F., Carpentier,M., Marcinkiewicz,M.,
Seidah,N.G., Crine,P., DesGroseillers,L. and Boileau,G.
TITLE     Molecular cloning and biochemical characterization of a new mouse
testis soluble zinc-metalloproteinase of the neprilysin family

```

Biochem. J. 347, 419-429. (2000)

2 (bases 1 to 2925)

Ghaddar, G., Ruchon, A.F., DesGroseillers, L. and Boileau, G.

Direct Submission

Submitted (09-AUG-1999) Biochemistry, University of Montreal, 2900

Edouard Montpetit, Montreal, Que H3T 1J4, Canada.

Location/Qualifiers

1..2925

/organism="Mus musculus"

/db_xref="taxon:10090"

/tissue_type="testis"

1..2925

/gene="Nli"

332..2629

/gene="Nli"

/product="neprilysin-like metalloproteinase 1"

/note="NLI; metallo-endopeptidase"

/codon_start=1

/protein_id="AAF69247.1"

/db_xref="GI:7769083"

/translation="MYERAGWGRKSPGVEGLMVLRLLLLGAIYVLGVFYSGIKQL
PLTLLSHFSWERTVVKALRRDSLSKIDCTPSCVIAARILENDSDRNCPCNFY
QACGCVLKRHDDEIPPLTNSRYSFVDILRDEVLTKGLVEDSTQHPAPVAKATLYRS
CMNQSVGWKRHSEPLSVLKVGWGPVAMDKNMTGLKWELEQLAVLNSQNRRLV
IDLFVNDPNSRRHVIYIDQPLGMPSEYVFOEDNNHKVRKAYLEFMTSVATMLRK
DONLSKESAMVREEMAEVLEETHLANATVPQEKRDVTALYHRMDLMELOERFGLK
FNTWTLFQNLVSLSEVELPDEVVYVIGLLEDDISYSARTMONTLWLRVLRLD
RIGLSORLEKREAVDYKALYGVTVVEVRWRECYSVNMSAVGSGLYIKRAFPSKLS
KTVRELIEKIRSFVNDLDELNMWDESKKAQERAMNIREQIGPYDITLEDNNKHL
DEEYSRLTFLPYFNGLOLNKNNAORSLKLREKVDQNLWIIIGAAVNAFYSPNRS
QIVFPAGILQPPFFSKDQPSLNFNGIGVMVIGHEITHGFDNDGRNPTDKNGMLDWSNS
FSARHFQQQSCQNCMIYQYGNFSEWELADONQVNGESTIGENIADNGVQYQAYKAVRLWLA
DGKKOORLPGNLNLITQAQFLFINTAQVWCGRSIRPEFAVQSITKTDVHSLPKTRVLGSLGN
LPGEFAEFKPRGSPMPMKRCRIR"

710 a 797 c 836 g 582 t

BASE COUNT.

ORIGIN

Query Match	72.6%	Score 237.4	DB 10	Length 2925
Best Local Similarity	82.9%	Pred. No. 5.8e-47		
Matches 271	Conservative	0	Mismatches 56	Indels 0
Gaps	0			
QY	1	ggcagcagatcatcagcagcgctttgatgcacaatgcccgaacttcgcacaagaatggcaac	60	
DB	2123	GGCAGCAGATCACACAGCGCTTTCATGATTAATGGTCTAACTTTTGACACAAAGCGCAAC	2182	
QY	61	atgatggattggtagagtaactctccaccagcaacttccggggcagctcagagtgcattg	120	
DB	2183	ATCTGGACTGGTGGAGTAACCTTCGCGCCGGCACCTTCCAACAGCAGTCGCAATGCATG	2242	
QY	121	atctaccagtacggcaactactctctggacctgcccagcagcagcaagaaactgaaacggattc	180	
DB	2243	ATCTATCAGTACGGCAACTTCTCTTTGGGAACATAGCAGCAACACCAAGAAATGTGAACGGATTTC	2302	
QY	181	aacaccttgggaaaaacattgctgcacaacgaggggtgcggcgaagcctataaagccctac	240	
DB	2303	AGTACCCCTCGGGGAGAACAATTGCCGACAACGGAGGTGTGCGACACGGCATACAAAGCTTAC	2362	
QY	241	ctcaagtggatggcagaggggtggcaagaccagcagctgcccgcgctggattctcaacccat	300	
DB	2363	CTACGGTGGCTGATGGGGCAAGATCAGCGACTGCCGGGACTCAACTGTACCTTAT	2422	
QY	301	gagcagctctcttcatacaactatgcc	327	
DB	2423	GCCAGCTTTTCTTCATCAACTATGCC	2449	

Search completed: October 6, 2002, 02:33:37
Job time: 3994 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2002, 01:27:58 ; Search time 211.84 Seconds
(without alignments)
2650.260 Million cell updates/sec

Title: US-09-647-780A-3
Perfect score: 327
Sequence: 1 gggcagagatcacgcacgg.....tcttcttcatacactatgcc 327

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues 3472872
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	325.4	99.5	2636	AAF59660	Human neprilysin-1
2	325.4	99.5	2663	AAF59661	Human neprilysin-1
3	325.4	99.5	2676	AA63764	CDNA encoding neut
4	325.4	99.5	2714	AAF59659	Human neprilysin-1
5	323.8	99.0	2076	AAF89737	Nucleotide sequenc
6	323.8	99.0	2262	AAF89739	Nucleotide sequenc
7	323.8	99.0	2340	AAF89738	Nucleotide sequenc
8	312.4	95.5	2332	AA597186	Human metalloprote
9	237.4	72.6	2925	AA63763	CDNA encoding neut

10	232.6	71.1	2765	20	AAZ28810	Rat membrane metal
11	203.4	62.2	1124	22	AAH26248	Human secreted pro
12	191	58.4	1226	22	AAH35007	Human colon cancer
13	159.6	48.8	5524	10	AAH90123	DNA encoding human
14	158	48.3	3181	9	AAH90952	Enkephalinase gene
15	158	48.3	3181	15	AAQ62678	Human enkephalinas
16	158	48.3	3181	22	AAH14514	Human neutral endo
17	158	48.3	3181	22	AAH14524	Human neutral endo
18	158	48.3	3181	22	AAH84073	Human enkephalinas
19	158	48.3	3181	22	AAH84083	Human enkephalinas
20	151	46.2	822	13	AAQ22244	Short-form rat enk
21	151	46.2	3243	15	AAH82261	Enkephalinase gene
22	151	46.2	3243	15	AAQ62679	Rat enkephalinase
23	151	46.2	3369	9	AAH81690	Atrial natriuretic
24	138.2	42.3	1986	23	ABL10163	Drosophila melanog
25	128.4	39.3	570	16	AAQ93868	Bovine endothelin
26	128.4	39.3	2129	16	AAQ93874	Bovine endothelin
27	128.4	39.3	2314	16	AAQ93860	Bovine endothelin
28	127.4	39.0	2889	19	AAH96149	Bovine endothelin
29	126.2	38.6	1703	16	AAQ93871	Bovine endothelin
30	120.4	36.8	2440	17	AAH42151	Human endothelin c
31	120.4	36.8	2440	17	AAH31399	Human endothelin c
32	120.4	36.8	2533	16	AAQ93877	Human placental en
33	120.4	36.8	2720	16	AAQ90315	Human ECE gene, H
34	120.4	36.8	2720	16	AAQ93861	Human endothelin c
35	117.8	36.0	2735	23	ABL15647	Drosophila melanog
36	111.8	34.2	1380	20	AAH34329	Human EST DNA49831
37	111.8	34.2	1380	21	AAH78595	Human EST DNA49831
38	111.8	34.2	1666	22	ABA06579	Human cDNA SEQ ID
39	111.8	34.2	2652	22	AAH80422	DNA encoding poly
40	110.2	33.7	2602	20	AAH34326	Human PRO403 nucle
41	110.2	33.7	2602	21	AAH78592	Human PRO403 nucle
42	110.2	33.7	2602	22	AAH46134	Human DNA encoding
43	110.2	33.2	3000	21	AAH08402	Human neurotransmi
44	108.6	33.2	7507	21	AAH38169	Metalloendoproteas
45	107	32.7	3291	19	AAV18976	Nucleotide sequenc

ALIGNMENTS

RESULT	1
AAF59660	
ID	AAF59660 standard; CDNA; 2636 BP.
XX	AAF59660;
AC	
XX	
DT	27-APR-2001 (first entry)
XX	
DE	Human neprilysin-like membrane metalloproteinase SNEPb CDNA.
XX	
KW	Human; SNEPb; neprilysin-like membrane metalloproteinase;
KW	splice variant; alternative splicing; zinc endopeptidase family;
KW	neuropeptide; peptide hormone; processing; metabolism; vaccine;
KW	drug screening; renal insufficiency; hepatic disease; ischaemia; pain;
KW	hypertension; cancer; inflammation; cardiovascular disease;
KW	neuronal disease; pancreatic disease; prostatic disease;
KW	respiratory disease; gene therapy; nephrotropic; vasotrophic; analgesic;
KW	cerebroprotective; hypotensive; cytostatic; antiinflammatory;
XX	cardiovascular; hepatotropic; ss.
OS	Homo sapiens.
XX	
PN	EP1069188-A1.
XX	
PD	17-JAN-2001.
XX	
PF	15-JUL-1999; 99EP-0401767.
XX	
PR	15-JUL-1999; 99EP-0401767.
XX	
PA	(SNFI) SANOFI-SYNTHELABO.
XX	

PI Jagerschmidt A, Agnel M, Culouscou J;
XX WPI; 2001-212582/22.
DR P-PSDB; AAB60562.
XX
XX
PT New membrane-associated metalloproteinase SNEPa, SNEPb and SNEPc.
PT polypeptides and polynucleotides, useful for treating e.g. acute and
PT chronic renal insufficiency, pain, stroke; cancer, inflammation, renal
PT and hepatic ischemia
XX
XX
PS Claim 5; Page 30-33; 72pp; English.
XX
XX The invention relates to the human SNEPa, SNEPb and SNEPc proteins, and
CC the cDNAs encoding them. SNEPa, SNEPb and SNEPc are neprilysin-like
CC membrane metalloproteinases and are the products of alternative splicing.
CC The substrate(s) for the SNEP proteins are not as yet known, although
CC the neprilysin family of zinc endopeptidases play key roles in the
CC processing and/or metabolism of neuro-peptides and peptide hormones. SNEP
CC nucleotides may be used as hybridisation probes for cDNA and genomic
CC DNA; to isolate full-length cDNAs and genomic clones encoding SNEPa,
CC SNEPb or SNEPc; to isolate cDNA and genomic clones of SNEP homologues;
CC as research reagents and material for the discovery of treatments and
CC diagnostics for animal and human diseases; and for chromosome
CC identification. The SNEP proteins may be used as immunogens to
CC produce antibodies immunospecific for SNEPa, SNEPb or SNEPc. Such
CC antibodies are used to isolate or identify clones expressing the
CC protein, or to purify the proteins by affinity chromatography.
XX
XX SNEP proteins may also be used in screening for compounds which modulate
CC SNEP endopeptidase activity, and to assess enzymatic cleavage of small
CC molecule substrates in cells, cell-free preparations, chemical libraries
CC and product mixtures. The SNEP proteins (as vaccine compositions),
CC to treat acute and chronic renal insufficiency, cancer, inflammation, as
CC ischaemia, pain, stroke, hypertensive disease, cancer, inflammation, as
CC well as cardiovascular, neuronal, pancreatic, prostatic, renal,
CC respiratory or hepatic diseases. They may also be used in modulating
CC peptide activation and/or degradation in the brain or kidney or in
CC another organ, or to diagnose or treat any disorder related to abnormal
CC expression of SNEPa, SNEPb or SNEPc. The present sequence represents
CC cDNA encoding SNEPb.
XX
XX Sequence 2636 BP; 601 A; 759 C; 813 G; 462 T; 1 other;
SQ
Query Match 99.5%; Score 325.4; DB 22; Length 2636;
Best Local Similarity 99.7%; Pred. No. 9.1e-81;
Matches 326; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 gggcagagatcacgcagcgtttgatgacaaatggccggaacttcgacagaatggcaac 60
Db 1763 gggcagagatcacgcagcgtttgatgacaaatggccggaacttcgacagaatggcaac 1822
Qy 61 atgatggattggtagtaacttctccaccgcaacttcgggagcagtcagagtgcacg 120
Db 1823 atgatggattggtagtaacttctccaccgcaacttcgggagcagtcagagtgcacg 1882
Qy 121 atctaccagtcgcaactactcctgggacacttcggagcagacagaaactgaacgattc 180
Db 1883 atctaccagtcgcaactactcctgggacacttcggagcagacagaaactgaacgattc 1942
Qy 181 aacacccctgggaaaaattctgacacagcaggggtgctggcgaagcctataaggcctac 240
Db 1943 aacacccctgggaaaaattctgacacagcaggggtgctggcgaagcctataaggcctac 2002
Qy 241 ctcaagtggatggcagaggggtggcgaagcagcagcgtgcccggcgtgattcaccctac 300
Db 2003 ctcaagtggatggcagaggggtggcgaagcagcagcgtgcccggcgtgattcaccctac 2062
Qy 301 gaggcagctctcttcatcaactatgcc 327
Db 2063 gaggcagctctcttcatcaactatgcc 2089

RESULT 2
AAF59661
ID AAF59661 standard; cDNA; 2663 BP.
XX
XX AC AAF59661;
XX
XX DT 27-APR-2001 (first entry)
XX
XX DE Human neprilysin-like membrane metalloproteinase SNEPc cDNA.
XX
XX KW Human; SNEPc; neprilysin-like membrane metalloproteinase;
KW splice variant; alternative splicing; zinc endopeptidase family;
KW neuro-peptide; peptide hormone; processing; metabolism; vaccine;
KW drug screening; renal insufficiency; hepatic disease; ischaemia; pain;
KW hypertension; cancer; inflammation; cardiovascular disease;
KW neuronal disease; pancreatic disease; prostatic disease;
KW respiratory disease; gene therapy; nephrotropic; vasotropic; analgesic;
KW cerebroprotective; hypotensive; cytostatic; antiinflammatory;
KW cardiovascular; hepatotropic; ss.
XX
XX OS Homo sapiens.
XX
XX PN EP1069188-A1.
XX
XX PD 17-JAN-2001.
XX
XX PF 15-JUL-1999; 99EP-0401767.
XX
XX PR 15-JUL-1999; 99EP-0401767.
XX
XX PA (SNFI) SANOFI-SYNTHELABO.
XX
XX PI Jagerschmidt A, Agnel M, Culouscou J;
XX
XX DR WPI; 2001-212582/22.
XX
XX DR P-PSDB; AAB60563.
XX
XX PT New membrane-associated metalloproteinase SNEPa, SNEPb and SNEPc.
XX polypeptides and polynucleotides, useful for treating e.g. acute and
XX chronic renal insufficiency, pain, stroke, cancer, inflammation, renal
XX and hepatic ischemia
XX
XX PS Claim 5; Page 35-38; 72pp; English.
XX
XX The invention relates to the human SNEPa, SNEPb and SNEPc proteins, and
CC the cDNAs encoding them. SNEPa, SNEPb and SNEPc are neprilysin-like
CC membrane metalloproteinases and are the products of alternative splicing.
CC The substrate(s) for the SNEP proteins are not as yet known, although
CC the neprilysin family of zinc endopeptidases play key roles in the
CC processing and/or metabolism of neuro-peptides and peptide hormones. SNEP
CC nucleotides may be used as hybridisation probes for cDNA and genomic
CC DNA; to isolate full-length cDNAs and genomic clones encoding SNEPa,
CC SNEPb or SNEPc; to isolate cDNA and genomic clones of SNEP homologues;
CC as research reagents and material for the discovery of treatments and
CC diagnostics for animal and human diseases; and for chromosome
CC identification. The SNEP proteins may be used as immunogens to
CC produce antibodies immunospecific for SNEPa, SNEPb or SNEPc. Such
CC antibodies are used to isolate or identify clones expressing the
CC protein, or to purify the proteins by affinity chromatography.
XX
XX SNEP proteins may also be used in screening for compounds which modulate
CC SNEP endopeptidase activity, and to assess enzymatic cleavage of small
CC molecule substrates in cells, cell-free preparations, chemical libraries
CC and product mixtures. The SNEP proteins (as vaccine compositions),
CC to treat acute and chronic renal insufficiency, cancer, inflammation, as
CC ischaemia, pain, stroke, hypertensive disease, cancer, inflammation, as
CC well as cardiovascular, neuronal, pancreatic, prostatic, renal,
CC respiratory or hepatic diseases. They may also be used in modulating
CC peptide activation and/or degradation in the brain or kidney or in
CC another organ, or to diagnose or treat any disorder related to abnormal
CC expression of SNEPa, SNEPb or SNEPc. The present sequence represents
CC cDNA encoding SNEPc.

SQ Sequence 2663 BP; 609 A; 769 C; 816 G; 468 T; 1 other;

Query Match 99.5%; Score 325.4; DB 22; Length 2663;
Best Local Similarity 99.7%; Pred. No. 9.1e-81;
Matches 326; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 gggcagagatcacgcacgagcttggatgacaaatggccggaacttcgacaaagaatggcaac 60
Db 1790 gggcagagatcacgcacgagcttggatgacaaatggccggaacttcgacaaagaatggcaac 1849
QY 61 atgatgattggtgagtaactctccaccagcaacttcctccgggagcagtcagagtgcatg 120
Db 1850 atgatgattggtgagtaactctccaccagcaacttcctccgggagcagtcagagtgcatg 1909
QY 121 atctaccagtcacggaactactctctggacctggcagacgacgacgaacgaactgaacggattc 180
Db 1910 atctaccagtcacggaactactctctggacctggcagacgacgacgaacgaactgaacggattc 1969
QY 181 aacaccttggggaaaacattgctgacaacgaggggtgctgcaagcctataaagcctac 240
Db 1970 aacaccttggggaaaacattgctgacaacgaggggtgctgcaagcctataaagcctac 2029
QY 241 ctcaagtggatggcagaggtggcaagaccagcagctgcccggcctggatctcacccat 300
Db 2030 ctcaagtggatggcagaggtggcaagaccagcagctgcccggcctggatctcacccat 2089
QY 301 gacgagctctcttctcaactatgccc 327
Db 2090 gacgagctctcttctcaactatgccc 2116

RESULT 3

AA63764

ID AAA63764 standard; cDNA; 2676 BP.

XX AC AAA63764;

XX 04-DEC-2000 (first entry)

XX cDNA: encoding neutral endopeptidase metalloproteinase-like enzyme NL-2.

XX Neprilysin; neutral endopeptidase metalloproteinase-like enzyme;

XX NEP-like enzyme; protein production; protein secretion;

XX neurological disease; Alzheimer's disease; pain; psychiatric disorder;

XX fertility; bone disease; abnormal phosphate metabolism; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 7..2319

XX /tag= a

XX /product= "neutral endopeptidase metalloproteinase-like

XX enzyme NL-2"

XX WO200047750-A2.

XX 17-AUG-2000.

XX 11-FEB-2000; 2000WO-CA00147.

XX 11-FEB-1999; 1999CA-2360376

XX (UIWO-) UNIV MONTREAL.

XX Desgroselliers L, Boileau G;

XX WPI: 2000-549148/50.

XX P-PSDB; AAB08131.

XX Novel neutral endopeptidase-like metalloproteinase polypeptides and

XX polynucleotides, used to screen for related sequences and enzyme

XX inhibitors, used for the treatment of NL-3 related bone disorders

XX PS

XX Disclosure; Fig 4; 59pp; English.

XX CC

XX The present sequence encodes a human neutral endopeptidase

XX metalloproteinase-like enzyme, designated NL-2. The specification

XX also describes NL-1 and NL-3. The NL enzymes are used to test for

XX specific inhibitors. The N-terminal region of the enzymes can be used

XX to promote production and secretion of foreign proteins and active

XX biopeptides, using chimeric constructs containing the foreign protein

XX downstream from and in phase with the N-terminal region. The NL enzymes

XX are have been localised to the brain, and may be useful in the

XX treatment of neurological diseases such as Alzheimer's disease, pain,

XX and psychiatric disorders. NL enzymes have also been localised to the

XX testis and ovaries, and may be used to control fertility. They have

XX also been localised to bone, and may be used to treat bone diseases,

XX and abnormal phosphate metabolisms related to improper peptide

XX processing by the NL-3 enzyme.

XX Sequence 2676 BP; 608 A; 771 C; 823 G; 474 T; 0 other;

SQ

Query Match 99.5%; Score 325.4; DB 21; Length 2676;

Best Local Similarity 99.7%; Pred. No. 9.1e-81;

Matches 326; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gggcagagatcacgcacgagcttggatgacaaatggccggaacttcgacaaagaatggcaac 60

Db 1813 gggcagagatcacgcacgagcttggatgacaaatggccggaacttcgacaaagaatggcaac 1872

QY 61 atgatgattggtgagtaactctccaccagcaacttcctccgggagcagtcagagtgcatg 120

Db 1873 atgatgattggtgagtaactctccaccagcaacttcctccgggagcagtcagagtgcatg 1932

QY 121 atctaccagtcacggaactactctctggacctggcagacgaggggtgctgcaagcctataaagcctac 180

Db 1933 atctaccagtcacggaactactctctggacctggcagacgaggggtgctgcaagcctataaagcctac 1992

QY 181 aacaccttggggaaaacattgctgacaacgaggggtgctgcaagcctataaagcctac 240

Db 1993 aacaccttggggaaaacattgctgacaacgaggggtgctgcaagcctataaagcctac 2052

QY 241 ctcaagtggatggcagaggtggcaagaccagcagctgcccggcctggatctcacccat 300

Db 2053 ctcaagtggatggcagaggtggcaagaccagcagctgcccggcctggatctcacccat 2112

QY 301 gacgagctctcttctcaactatgccc 327

Db 2113 gacgagctctcttctcaactatgccc 2139

RESULT 4

AAF59659

ID AAF59659 standard; cDNA; 2714 BP.

XX AC AAF59659;

XX 27-APR-2001 (first entry)

XX Human neprilysin-like membrane metalloproteinase SNEPA cDNA.

XX Human; SNEPA; neprilysin-like membrane metalloproteinase;

XX splice variant; alternative splicing; zinc endopeptidase family;

XX neuropeptide; peptide hormone; processing; metabolism; vaccine;

XX drug screening; renal insufficiency; hepatic disease; ischaemia; pain;

XX hypertension; cancer; inflammation; cardiovascular disease;

XX neuronal disease; pancreatic disease; prostatic disease;

XX respiratory disease; gene therapy; nephrotropic; vasotropic; analgesic;

XX cerebroprotective; hypotensive; cytostatic; antiinflammatory;

XX cardiovascular; hepatotropic; ss.

XX Homo sapiens.

XX EPI069188-A1.

PN

xx PD 17-JAN-2001.
 xx PF 15-JUL-1999; 99EP-0401767.
 xx PR 15-JUL-1999; 99EP-0401767.
 xx PA (SNFI) SANOFI-SYNTHELABO.
 xx PI Jagerschmidt A, Agnel M, Culouscou J;
 xx DR WPI; 2001-212582/22.
 xx DR P-PSDB; AAB60561.
 xx
 xx New membrane-associated metalloproteinase SNEPa, SNEPb and SNEPc.
 xx PT polypeptides and polynucleotides, useful for treating e.g. acute and
 xx PT chronic renal insufficiency, pain, stroke, cancer, inflammation, renal
 xx PT and hepatic ischemia
 xx PS Claim 5; Page 25-28; 72pp; English.
 xx
 xx The invention relates to the human SNEPa, SNEPb and SNEPc proteins, and
 xx CC the cDNAs encoding them. SNEPa, SNEPb and SNEPc are neprilysin-like
 xx CC membrane metalloproteinases and are the products of alternative splicing.
 xx CC The substrate(s) for the SNEP proteins are not as yet known, although
 xx CC the neprilysin family of zinc endopeptidases play key roles in the
 xx CC processing and/or metabolism of neuro-peptides and peptide hormones. SNEP
 xx CC nucleotides may be used as hybridisation probes for cDNA and genomic
 xx CC DNA; to isolate full-length cDNAs and genomic clones encoding SNEPa,
 xx CC SNEPb or SNEPc; to isolate cDNA and genomic clones of SNEP homologues;
 xx CC as research reagents and material for the discovery of treatments and
 xx CC diagnostics for animal and human diseases; and for chromosome
 xx CC identification. The SNEP proteins may be used as immunogens to
 xx CC produce antibodies immunospecific for SNEPa, SNEPb or SNEPc. Such
 xx CC antibodies are used to isolate or identify clones expressing the
 xx CC protein, or to purify the proteins by affinity chromatography.
 xx CC SNEP proteins may also be used in screening for compounds which modulate
 xx CC SNEP endopeptidase activity, and to assess enzymatic cleavage of small
 xx CC molecule substrates in cells, cell-free preparations, chemical libraries
 xx CC and product mixtures. The SNEP proteins (as vaccine compositions),
 xx CC SNEP nucleotides, and SNEP activators or inhibitors may be used
 xx CC to treat acute and chronic renal insufficiency, cancer, inflammation, as
 xx CC ischaemia, pain, stroke, hypertensive disease, pancreatic, renal,
 xx CC well as cardiovascular, neuronal, pancreatic, prostatic, renal,
 xx CC respiratory or hepatic diseases. They may also be used in modulating
 xx CC peptide activation and/or degradation in the brain or kidney or in
 xx CC another organ, or to diagnose or treat any disorder related to abnormal
 xx CC expression of SNEPa, SNEPb or SNEPc. The present sequence represents
 xx CC cDNA encoding SNEPa.
 xx SQ Sequence 2714 BP; 621 A; 780 C; 836 G; 476 T; 1 other;

Query Match 99.5%; Score 325.4; DB 22; Length 2714;
 Best Local Similarity 99.7%; Pred. No. 9.2e-81;
 Matches 326; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gggcagagatcagcagcgggttgatgacaaatggccggaaacttcgacaaagaatggcaac 60
 |||
 Db 1841 gggcagagatcagcagcgggttgacagaaatggccggaaacttcgacaaagaatggcaac 1900
 |||
 Qy 61 atgatgattggtgagtaactctccaccagcaacttcgggagcagtcagatgcatg 120
 |||
 Db 1901 atgatgattggtgagtaactctccaccagcaacttcgggagcagtcagatgcatg 1960
 |||
 Qy 121 atctaccagtcagcagcaactctctgggacctggcagcagcaacagcgtgaacgattc 180
 |||
 Db 1961 atctaccagtcagcagcaactctctgggacctggcagcagcaacagcgtgaacgattc 2020
 |||
 Qy 181 aacaccttggggaaaaacattctgacaaacgaggggttcgggcaagcctataagcctac 240
 |||
 Db 2021 aacaccttggggaaaaacattctgacaaacgaggggttcgggcaagcctataagcctac 2080
 |||

Qy 241 ctcaagtggatggcagaggggtggcaagaccagcagctgccggcctggatctccaccat 300
 |||
 Db 2081 ctcaagtggatggcagaggggtggcaagaccagcagctgccggcctggatctccaccat 2140
 |||
 Qy 301 gagcagctctcttcttcatacaactatgcc 327
 |||
 Db 2141 gagcagctctcttcttcatacaactatgcc 2167
 |||

RESULT 5
 AAF89737
 ID AAF89737 standard; DNA; 2076.BP.
 XX
 AC AAF89737;
 XX
 DT 23-JUL-2001 (first entry)
 XX
 DE Nucleotide sequence of a human metalloprotease enzyme IGS5.
 XX
 KW Metalloprotease; IGS5; infection; pain; cancer; diabetes; obesity;
 KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;
 KW hypotension; hypertension; urinary retention; osteoporosis;
 KW angina pectoris; myocardial infarction; stroke; ulcer; allergy;
 KW benign prostatic hypertrophy; migraine; psychotic disorder;
 KW neurological disorder; autism; multiple sclerosis; Alzheimer's disease;
 KW neurodegenerative disease; sleep disorder; epilepsy; kidney disease;
 KW cardiovascular disease; arteriosclerosis; cerebrovasospasm;
 KW subarachnoid hemorrhage; cerebral ischemia; cerebral infarction;
 KW peripheral vascular disease; Raynaud's disease; motility disorder;
 KW gastrointestinal disorder; gastric emptying; gastroparesis; diarrhoea;
 KW inflammation; chemotherapy induced injury; tumour invasion;
 KW immune disorder; arthritis; endotoxin shock; sepsis; diabetes mellitus;
 KW severe mental retardation; dyskinesia; Huntington's disease;
 KW Gilles de la Tourette's syndrome; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2076
 FT /*tag= a
 FT /product= "metalloprotease enzyme IGS5"
 FT
 FT WO200136610-A1.
 PD 25-MAY-2001.
 XX
 XX 17-NOV-2000; 2000MO-EPI1532.
 XX
 XX 19-NOV-1999; 99EP-0203862.
 XX 19-NOV-1999; 99NL-1013616.
 XX 31-MAY-2000; 2000EP-0201937.
 XX 31-MAY-2000; 2000NL-1015356.
 XX
 XX (SOLV) SOLVAY PHARM BV.
 XX
 XX Deleersnijder W, Wiegers R, Weske M;
 XX
 XX WPI; 2001-343815/36.
 XX P-PSDB; AAB83840.
 XX
 XX New IGS5 polypeptides useful for treating infections, pain, cancer,
 XX diabetes, obesity, anorexia, bulimia, asthma, schizophrenia,
 XX hypertension, urinary retention and Parkinson's disease
 XX
 XX Claim 11; Page 5-6; 115pp; English.
 XX
 XX The present sequence encodes a human metalloprotease enzyme designated
 XX IGS5. IGS5 polynucleotides and polypeptides are useful for treating
 XX infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma,
 XX Parkinson's disease, acute heart failure, hypotension, hypertension,
 XX urinary retention, osteoporosis, angina pectoris, myocardial infarction,
 XX stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,
 XX psychotic and neurological disorders, autism, multiple sclerosis,

CC Alzheimer's disease, and other neurodegenerative diseases, sleep
 CC disorder, epilepsy, cardiovascular diseases, arteriosclerosis,
 CC cerebrovasospasm, subarachnoid hemorrhage, cerebral ischemia, cerebral
 CC infarction, peripheral vascular disease, Raynaud's disease, kidney
 CC diseases, gastrointestinal disorders, motility disorders and conditions
 CC of delayed gastric emptying, post-operative or diabetic gastroparesis,
 CC diarrhoea, inflammations, chemotherapy induced injury, tumour invasion,
 CC immune disorders, arthritis, endotoxin shock, sepsis, complications of
 CC diabetes mellitus, and severe mental retardation and dyskinesias, such
 CC as Huntington's disease or Gilles de la Tourette's syndrome.

XX
 SQ Sequence 2076 BP; 493 A; 578 C; 636 G; 369 T; 0 other;

Query Match 99.0%; Score 323.8; DB 22; Length 2076;
 Best Local Similarity 99.4%; Pred. No. 2.3e-80;

Matches 325; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggcagagatcacgcagcgtttgatgacaatggccggaacttcgacaagaatggcaac 60
 DB 1570 gggcagagatcacgcagcgtttgatgacaatggccggaacttcgacaagaatggcaac 1629

QY 61 atgatgattggtagtaactctccaccagcactccgggagcagtcagatgcatg 120
 DB 1630 atgatgattggtagtaactctccaccagcactccgggagcagtcagatgcatg 1689

QY 121 atctaccagtcagcgaactactctctgggaactggcagcagacgaacgtgaacggattc 180
 DB 1690 atctaccagtcagcgaactactctctgggaactggcagcagacgaacgtgaacggattc 1749

QY 181 aacaccccttgggaaaaattctgacaacgaggggtggcgaagcctataagcctac 240
 DB 1750 aacaccccttgggaaaaattctgacaacgaggggtggcgaagcctataagcctac 1809

QY 241 ctcaagtggatggcagaggggtggcaagaccagcagctgcccgccgtgattccaccat 300
 DB 1810 ctcaagtggatggcagaggggtggcaagaccagcagctgcccgccgtgattccaccat 1869

QY 301 gaggcagctcttcttcatacaactatgcc 327
 DB 1870 gaggcagctcttcttcatacaactatgcc 1896

RESULT 6

AAF89739
 ID AAF89739 standard; DNA; 2262 BP.

XX
 AC AAF89739;

DT 23-JUL-2001 (first entry)

XX Nucleotide sequence of a human metalloprotease enzyme IGS5.

XX Metalloprotease; IGS5; infection; pain; cancer; diabetes; obesity;
 KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;
 KW hypotension; hypertension; urinary retention; osteoporosis;
 KW angina pectoris; myocardial infarction; stroke; ulcer; allergy;
 KW benign prostatic hypertrophy; migraine; psychotic disorder;
 KW neurological disorder; autism; multiple sclerosis; Alzheimer's disease;
 KW neurodegenerative disease; sleep disorder; epilepsy; kidney disease;
 KW cardiovascular disease; arteriosclerosis; cerebrovasospasm;
 KW subarachnoid hemorrhage; cerebral ischemia; cerebral infarction;
 KW peripheral vascular disease; Raynaud's disease; motility disorder;
 KW gastrointestinal disorder; gastric emptying; gastroparesis; diarrhoea;
 KW inflammation; chemotherapy induced injury; tumour invasion;
 KW immune disorder; arthritis; endotoxin shock; sepsis; diabetes mellitus;
 KW severe mental retardation; dyskinesia; Huntington's disease;
 KW Gilles de la Tourette's syndrome; ss.

XX Homo sapiens.

OS
 XX
 XX
 FH Location/Qualifiers
 FT CDS 1..2262

FT
 FT
 XX
 PN
 XX
 XX
 PD
 XX
 PF
 XX
 PR
 PR
 PR
 PR
 XX
 PA
 PI
 PI
 DR
 DR
 XX
 PT
 PT
 PT
 XX
 PS
 XX
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 SQ

W0200136610-A1.

25-MAY-2001.

17-NOV-2000; 2000WO-EP11532.

19-NOV-1999; 99EP-0203862.

31-MAY-2000; 2000EP-0201937.

31-MAY-2000; 2000NL-1015356.

(SOLV) SOLVAY PHARM BV.

Deleersnijder W, Wieggers R, Weske M;

WPI; 2001-343815/36.

P-PSDB; AAB83842.

New IGS5 polypeptides useful for treating infections, pain, cancer,
 diabetes, obesity, anorexia, bulimia, asthma, schizophrenia,
 hypertension, urinary retention and Parkinson's disease

Claim 11; Page 8-9; 115pp; English.

The present sequence encodes a human metalloprotease enzyme designated
 IGS5. IGS5 polynucleotides and polypeptides are useful for treating
 infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma,
 Parkinson's disease, acute heart failure, hypotension, hypertension,
 urinary retention, osteoporosis, angina pectoris, myocardial infarction,
 stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,
 psychotic and neurological disorders, autism, multiple sclerosis,
 Alzheimer's disease, and other neurodegenerative diseases, sleep
 disorder, epilepsy, cardiovascular diseases, arteriosclerosis, cerebral
 cerebrovasospasm, subarachnoid hemorrhage, cerebral ischemia, cerebral
 infarction, peripheral vascular disease, Raynaud's disease, kidney
 diseases, gastrointestinal disorders, motility disorders and conditions
 of delayed gastric emptying, post-operative or diabetic gastroparesis,
 diarrhoea, inflammations, chemotherapy induced injury, tumour invasion,
 immune disorders, arthritis, endotoxin shock, sepsis, complications of
 diabetes mellitus, and severe mental retardation and dyskinesias, such
 as Huntington's disease or Gilles de la Tourette's syndrome.

Sequence 2262 BP; 520 A; 628 C; 716 G; 398 T; 0 other;

Query Match 99.0%; Score 323.8; DB 22; Length 2262;

Best Local Similarity 99.4%; Pred. No. 2.4e-80;

Matches 325; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggcagagatcacgcagcgtttgatgacaatggccggaacttcgacaagaatggcaac 60

DB 1756 gggcagagatcacgcagcgtttgatgacaatggccggaacttcgacaagaatggcaac 1815

QY 61 atgatgattggtagtaactctccaccagcactccgggagcagtcagatgcatg 120

DB 1816 atgatgattggtagtaactctccaccagcactccgggagcagtcagatgcatg 1875

QY 121 atctaccagtcagcgaactactctctgggaactggcagcagacgaacgtgaacggattc 180

DB 1876 atctaccagtcagcgaactactctctgggaactggcagcagacgaacgtgaacggattc 1935

QY 181 aacaccccttgggaaaaattctgacaacgaggggtggcgaagcctataagcctac 240

DB 1936 aacaccccttgggaaaaattctgacaacgaggggtggcgaagcctataagcctac 1995

QY 241 ctcaagtggatggcagaggggtggcaagaccagcagctgcccgccgtgattccaccat 300

DB 1996 ctcaagtggatggcagaggggtggcaagaccagcagctgcccgccgtgattccaccat 2055

QY 301 gaggcagctcttcttcatacaactatgcc 327

DB 2056 gagcagctcttctcactcaactacgcc 2082
|||||

RESULT 7

AAF89738
ID AAF89738 standard; DNA; 2340 BP.

XX AC AAF89738;

XX DT 23-JUL-2001 (first entry)

XX DE Nucleotide sequence of a human metalloprotease enzyme IGS5.

XX KW Metalloprotease; IGS5; infection; pain; cancer; diabetes; obesity;
XX KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;
XX KW hypotension; hypertension; urinary retention; osteoporosis;
XX KW angina pectoris; myocardial infarction; stroke; ulcer; allergy;
XX KW benign prostatic hypertrophy; migraine; psychotic disorder;
XX KW neurological disorder; autism; multiple sclerosis; Alzheimer's disease;
XX KW neurodegenerative disease; sleep disorder; epilepsy; kidney disease;
XX KW cardiovascular disease; arteriosclerosis; cerebrovasospasm;
XX KW subarachnoid hemorrhage; cerebral ischemia; cerebral infarction;
XX KW peripheral vascular disease; Raynaud's disease; motility disorder;
XX KW gastrointestinal disorder; gastric emptying; gastroparesis; diarrhoea;
XX KW inflammation; chemotherapy induced injury; tumour invasion;
XX KW immune disorder; arthritis; endotoxin shock; sepsis; diabetes mellitus;
XX KW severe mental retardation; dyskinesia; Huntington's disease;
XX KW Gilles de la Tourette's syndrome; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 1..2340

XX FT /tag= a

XX FT /product= "metalloprotease enzyme IGS5"

XX PN WO200136610-A1.

XX PD 25-MAY-2001.

XX PF 17-NOV-2000; 2000WO-EP11532.

XX PR 19-NOV-1999; 99EP-0203862.

XX PR 19-NOV-1999; 99NL-1013616.

XX PR 31-MAY-2000; 2000EP-0201937.

XX PR 31-MAY-2000; 2000NL-1015356.

XX PA (SOLV) SOLVAY PHARM BV.

XX PI Deleersnijder W, Wiegers R, Weske M;

XX XX WPI; 2001-343815/36.

XX DR P-PSDB; AAB83841.

XX PT New IGS5 polypeptides useful for treating infections, pain; cancer,
XX PT diabetes, obesity, anorexia, bulimia, asthma, schizophrenia,
XX PT hypertension, urinary retention and Parkinson's disease

XX PS Claim 11; Page 6-7; 115pp; English.

XX CC The present sequence encodes a human metalloprotease enzyme designated
XX CC IGS5. IGS5 polynucleotides and polypeptides are useful for treating
XX CC infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma,
XX CC Parkinson's disease, acute heart failure, hypotension, hypertension,
XX CC urinary retention, osteoporosis, angina pectoris, myocardial infarction,
XX CC stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,
XX CC psychotic and neurological disorders, autism, multiple sclerosis,
XX CC Alzheimer's disease, and other neurodegenerative diseases, sleep
XX CC disorder, epilepsy, cardiovascular diseases, arteriosclerosis,
XX CC cerebrovasospasm, subarachnoid hemorrhage, cerebral ischemia, cerebral
XX CC infarction, peripheral vascular disease, Raynaud's disease, kidney
XX CC diseases, gastrointestinal disorders, motility disorders and conditions

CC of delayed gastric emptying, post-operative or diabetic gastroparesis,
CC diarrhoea, inflammations, chemotherapy induced injury, tumour invasion,
CC immune disorders, arthritis, endotoxin shock, sepsis, complications of
CC diabetes mellitus, and severe mental retardation and dyskinesias, such
CC as Huntington's disease or Gilles de la Tourette's syndrome.

XX SQ Sequence 2340 BP; 539 A; 649 C; 739 G; 413 T; 0 other;

Query Match 99.08; Score 323.8; DB 22; Length 2340;
Best Local Similarity 99.4%; Pred. No. 2.4e-80;
Matches 325; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gggcagcagatcacgcgcgctttgatgacaaatggccggaacttcgacaaagaatggcaac 60
|||||
DB 1834 gggcagcagatcacgcgcgctttgatgacaaatggccggaacttcgacaaagaatggcaac 1893
|||||

QY 61 atgatgattggtggagtaacttctccaccagcacttcgaggagcagtcagatgcatg 120
|||||
DB 1894 atgatgattggtggagtaacttctccaccagcacttcgaggagcagtcagatgcatg 1953
|||||

QY 121 atctaccagtagcgaactactctctggacctggcagcagcagacagacgtgaacgattc 180
|||||
DB 1954 atctaccagtagcgaactactctctggacctggcagcagcagacagacgtgaacgattc 2013
|||||

QY 181 aacaccttgggggaaaacattgctgacaaacgagggtgctgcaagcctataagcctac 240
|||||
DB 2014 aacaccttgggggaaaacattgctgacaaacgagggtgctgcaagcctataagcctac 2073
|||||

QY 241 ctcaagtggatggcagaggggtggcaagaccagcagctgcccgctggatctcaccat 300
|||||
DB 2074 ctcaagtggatggcagaggggtggcaagaccagcagctgcccgctggatctcaccat 2133
|||||

QY 301 gagcagctcttcttcatacaactatgcc 327
|||||

DB 2134 gagcagctcttcttcatacaactatgcc 2160

RESULT 8

AA597186

ID AA597186 standard; cDNA; 2232 BP.

XX AC AA597186;

XX DT 26-FEB-2002 (first entry)

XX DE Human metalloprotease partial DNA sequence #15.

XX KW Human; protease; PCR primer; cytostatic; immunomodulator; cardiant;
XX KW vasotropic; antimigraine; analgesic; endocrine; neurotropic; tranquiliser;
XX KW hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic;
XX KW anorectic; antiinflammatory; aspartyl protease; cysteine protease;
XX KW metalloprotease; serine protease; cancer; haematopoietic; breast; colon;
XX KW lung; prostate; cervical; brain; ovarian; bladder; kidney; pain;
XX KW immune-related disease; cardiovascular disease; neuronal disease;
XX KW migraine; sexual dysfunction; mood disorder; attention disorder;
XX KW cognition disorder; hypotension; hypertension; psychotic disorder;
XX KW dyskinesia; metabolic disorder; inflammatory disorder; ss.

XX OS Homo sapiens.

XX PN WO200183782-A2.

XX PD 08-NOV-2001.

XX PF 04-MAY-2001; 2001WO-US14431.

XX PR 04-MAY-2000; 2000US-201879P.

XX PA (SUGE-) SUGEN INC.

XX PI Plowman GP, Whyte D, Sudarsanam S, Manning G, Caenepeel S;

XX PI Payne V;

```
XX WPI; 2002-041502/05.
DR P-PSDB; AAU72903.
XX
XX Novel protease polypeptide useful for screening for substances that may
PT be used to treat, e.g., cancers, immune-related diseases,
PT cardiovascular disease, migraine, pain, psychotic and inflammatory
PT disorders.
XX
XX Claim 30; Figure 1R-S; 232pp; English.
PS
XX The invention relates to an isolated, enriched, or purified protease
CC polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to
CC screen for substances (S) that may modulate its activity. Administering
CC S (which modulates protease activity in vitro) may be used to treat a
CC disease or disorder selected from cancers (e.g., of tissues, of blood or
CC haematopoietic origin, of the breast, colon, lung, prostate, cervical,
CC brain, ovarian, bladder or kidney), immune-related diseases and
CC disorders, cardiovascular disease, brain or neuronal-associated diseases
CC (e.g., central or peripheral nervous system diseases, migraine, pain,
CC sexual dysfunction, mood disorders, attention disorders, cognition
CC disorders, hypotension, hypertension, psychotic disorders, neurological
CC disorders and dyskinesias), metabolic disorders and inflammatory
CC disorders. (I) may also be useful as a diagnostic tool for a disease or
CC disorder such as those above. AAS97159-AAS97195 represent human
CC protease coding sequences and primers of the invention.
XX
XX Sequence 2232 BP; 512 A; 620 C; 705 G; 395 T; 0 other;
SQ
Query Match 95.5%; Score 312.4; DB 24; Length 2232;
Best Local Similarity 98.8%; Pred. No. 3 6e-77;
Matches 326; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
QY 1 gggcagagatcacgcacgggtttgatgacaatg---gccggaacttcgacaaagaatgccc 57
Db 1723 gggcagagatcacgcacgggtttgatgacaatggtggccggaacttcgacaaagaatgccc 1782
QY 58 aacatgatgattggtgagtaactttccaccagacacttcctcgggagcagtcagagatgccc 117
Db 1783 aacatgatgattggtgagtaactttccaccagacacttcctcgggagcagtcagagatgccc 1842
QY 118 atgatctaccagtcacgcaactactctctgggacacttggcagacgacagacagtcgtaacgga 177
Db 1843 atgatctaccagtcacgcaactactctctgggacacttggcagacgacagacagtcgtaacgga 1902
QY 178 tcaacacccttggggaaacattgctgacaacgagaggggtgcggaagcctataagccc 237
Db 1903 tcaacacccttggggaaacattgctgacaacgagaggggtgcggaagcctataagccc 1962
QY 238 tactcgaagtggatggcagaggggtggaagacagcagtcgcccggcctggatctacc 297
Db 1963 tactcgaagtggatggcagaggggtggaagacagcagtcgcccggcctggatctacc 2022
QY 298 catgagcagctctcttcatacaactatgcc 327
Db 2023 catgagcagctctcttcatacaactatgcc 2052
RESULT 9
ID AAA63763
AC AAA63763 standard; cDNA; 2925 BP.
XX
XX AAA63763;
XX
XX 04-DEC-2000 (first entry)
XX
XX cDNA encoding neutral endopeptidase metalloproteinase-like enzyme, NL-1.
XX Neprilysin; neutral endopeptidase metalloproteinase-like enzyme;
KW NEP-like enzyme; protein production; protein secretion;
KW neurological disease; Alzheimer's disease; pain; psychiatric disorder;
KW fertility; bone disease; abnormal phosphate metabolism; ss.
XX
```

```
XX Mus sp.
OS
XX
XX Key Location/Qualifiers
FH CDS 332..2629
FT /*tag= a
FT /product= "neutral endopeptidase metalloproteinase-like
ET enzyme NL-1"
XX
XX WO200047750-A2.
PN
XX
XX 17-AUG-2000.
PD
XX
XX 11-FEB-2000; 2000WO-CA00147.
PF
XX
XX 11-FEB-1999; 99CA-2260376.
PR
XX (UYMO-) UNIV MONTREAL.
PA
XX Desgroseillers L, Bolleau G;
PI
XX WPI; 2000-549148/50.
DR P-PSDB; AAB08130.
XX
XX Novel neutral endopeptidase-like metalloproteinase polypeptides and
PT polynucleotides, used to screen for related sequences and enzyme
PT inhibitors, used for the treatment of NL-3 related bone disorders
XX
XX Disclosure; Fig 3; 59pp; English.
XX
XX The present sequence encodes a murine neutral endopeptidase
CC metalloproteinase-like enzyme, designated NL-1. The specification
CC also describes NL-2 and NL-3. The NL enzymes are used to test for
CC specific inhibitors. The N-terminal region of the enzymes can be used
CC to promote production and secretion of foreign proteins and active
CC bioproteins, using chimeric constructs containing the foreign protein
CC downstream from and in phase with the N-terminal region. The NL enzymes
CC have been localised to the brain, and may be useful in the
CC treatment of neurological diseases such as Alzheimer's disease, pain,
CC and psychiatric disorders. NL enzymes have also been localised to the
CC testis and ovaries, and may be used to control fertility. They have
CC also been localised to bones, and may be used to treat bone diseases,
CC and abnormal phosphate metabolisms related to improper peptide
CC processing by the NL-3 enzyme.
XX
XX Sequence 2925 BP; 710 A; 797 C; 836 G; 582 T; 0 other;
SQ
Query Match 72.6%; Score 237.4; DB 21; Length 2925;
Best Local Similarity 82.9%; Pred. No. 2 8e-56;
Matches 271; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 1 gggcagcagatcacgcacgggtttgatgacaacttcggaacttcgacaaagaatgccc 60
Db 2123 gggcagcagatcacgcacgggtttgatgacaacttcggaacttcgacaaagaatgccc 2182
QY 61 atgatggattggtgagtaacttctccaccagcacttcggaagcagtcagagtcagtc 120
Db 2183 atgctggactggtgagtaacttctccaccagcacttcggaagcagtcagagtcagtc 2242
QY 121 atctaccagtcagcgaactactctctgggacacttcggaagcagtcagagtcagtc 180
Db 2243 atctaccagtcagcgaactactctctgggacacttcggaagcagtcagagtcagtc 2302
QY 181 aacacccttggggaaacattgctgacaacgagaggggtgcggaagcctataagccc 240
Db 2303 agtacccttggggaaacattgctgacaacgagaggggtgcggaagcctataagccc 2362
QY 241 ctcaagtcagtcagcagaggggtggaagacagcagtcgcccggcctggatctacc 300
Db 2363 ctcaagtcagtcagcagaggggtggaagacagcagtcgcccggcctggatctacc 2422
QY 301 gacgagctctcttcatacaactatgcc 327
```

181 aacacccttgggaaaaacatttactgacaaacggagggtacggcaagcctataaacgcctac:240

AAH32943 to AAH37195 and AAG77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytostatic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing P. Inactive proteins or to supplement the patients own production of P. Additionally, N may be used to produce the colon cancer-associated Ps, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent sequences used in the exemplification of the present invention.

N.B. Pages 666 to 682 and page.7053 of the sequence listing were missing at time of publication, meaning no sequences are present for CC SEQ ID NO:1027 to 1052, 7921 and 7922.

Sequence.1226 BP; 270 A; 388 C; 345 G; 213 T; 110 other:

Query Match	58.4%;	Score 191;	DB 22;	Length 1226;
Best Local Similarity	73.7%;	Pred. No. 1.8e-43;		
Matches 323;	Conservative 1;	Mismatches 2;	Indels 112;	Gaps

QY	1	gggcacgagatca	cgacagcgctt	gatgacaatg	cgcggaacttc	gacaagaatg	ggcaac	60
Db	140	gggcagagatca	cgacagcgctt	nacgacaatg	cgcggaacttc	gacaagaatg	ggcaac	199
QY	61	atgatggattg	tgagtagtaact	tccaccagcact	tcgggagcag	tcagagtgcatg		120
Db	200	atgatggattg	tgagtagtaact	tccaccagcact	tcgggagcag	tcagagtgcatg		259
QY	121	atataccagtc	agcgaaactact	ctctggacctg	gcgacgaa	cagaac		168
Db	260	atataccagtc	agcgaaactact	ctctggga	-ctggcagac	aacagactg	agcgctgcc	318
QY	169	-----	-----	-----	-----	-----	-----	168
Db	319	accagacccag	gtcggggtacc	cgagagcccy	agccctg	gcctgagggagag	gggaag	378
QY	169	-----	-----	-----	-----	gtgaacg	gattcaac	189
Db	379	tcaggccggg	ggtgcgcccaact	ctctctctg	tgcgcagtg	aacggattca	acacccct	438
QY	190	ggggaaaaact	tgtcaaacgag	gggtgcgcg	caagcctata	gaagcctac	ctcaagtg	249
Db	439	ggggaaaaact	tgtcaaacgag	gggggtgcgcg	caagcctata	gaagcctac	ctcaagtg	498
QY	250	atggcagaggt	gtggcaagc	acgagcagctg	cccgcgctg	gatctc	accatgag	309
Db	499	atggcagaggt	gtggcaagc	acgagcagctg	cccgcgctg	gatctc	accatgag	558
QY	310	ttcttcatca	actatgcc	327				
Db	559	ttcttcatca	actatgcc	576				

RESULT	13
AAN90123	
ID	AAN90123 standard; DNA; 5524 BP.
XX	
AC	
AAN90123;	
XX	
01-NOV-1989	(first entry)
DT	
DE	DNA encoding human common acute lymphoblastic leukaemia antigen.
XX	
KW	DNA; human common acute lymphoblastic leukaemia antigen;
KW	analgesic; inhibits leukaemia; endopeptidase.

Db 1983 cagaattattataaaagatggcgaagaaataattacttccctggacttgacctaataatcac 2042
 QY 301 gggcagctcttcttcaactatgc 326
 Db 2043 aaacaactattttcttgaaacttgc 2068

RESULT 15

AA062678
 ID AA062678 standard; cDNA; 3181 BP.
 AC AA062678;
 XX
 DT 05-JAN-1995 (first entry)
 XX
 DE Human enkephalinase cDNA.
 XX
 KW Human; enkephalinase; neutral endopeptidase; kidney; intestine; kinin;
 KW kidney brush border neutral proteinase; pituitary; brain; lymph nodes;
 KW -neutrophils; enkephalins; tachykinins; ss.
 XX
 OS Homo sapiens.

Location/Qualifiers

Key 3..2249
 CDS /tag= a
 FT mat_peptide 21..2246
 FT /tag= b
 FT allele 1413
 FT /tag= c
 FT /note= "Nucleotide is A in a different clone
 changing the codon to Thr"

PN EP596355-A.

PD 11-MAY-1994.

PF 23-DEC-1987; 87EP-0117230.

PR 24-DEC-1986; 86US-0946566.

PR 12-JAN-1987; 87US-0002478.

XX (GETH) GENENTECH INC.

XX Malfroy-Camine B, Schofield PR;

XX WPI: 1994-152785/19.

XX P-PSDB; AAR52706.

XX Method for assay of enkephalinase using dansylated peptide(s) -
 for identification of recombinant enkephalinase

XX Disclosure; Fig 1; 61pp; English.

XX This sequence encodes human enkephalinase. Enkephalinase is a
 CC neutral endopeptidase or kidney brush border neutral proteinase. It
 CC has been isolated from the kidney, intestine, pituitary, brain and
 CC lymph nodes, and has been detected in many peripheral organs and in
 CC human neutrophils. The distribution of enkephalinase in the brain
 CC closely parallels the distribution of the enkephalins. Mammalian
 CC enkephalinase may be used in the treatment of various pathological
 CC disorders associated with various endogenous peptides, eg.
 CC tachykinins and kinins.

XX Sequence 3181 BP; 1055 A; 582 C; 657 G; 887 T; 0 other;

Query Match 48.3%; Score 158; DB 15; Length 3181;
 Best Local Similarity 67.8%; Pred. No. 3.6e-34;
 Matches 221; Conservative 0; Mismatches 105; Indels 0; Caps 0;

QY 1 gggcagatcacgcagcggctttgatgacaatggcgggaacttcgacaagaatggcaac 60

Db 1743 ggacacgaaatcacccatggcttcgatgacaatggcagaactttaacaaagatggagac 1802
 QY 61 atgatggattgggtggagtaactttccaccacacacttcgggagcagtcagagtgcatg 120
 Db 1803 ctgcttgactgggtggactcaacagtcgtcaagtaactttaaggagcaatcccatgcatg 1862
 QY 121 atctaccagtacggcaactactctctgggacctggcagacgaacagaacgtgaacggattc 180
 Db 1863 gtgatcagatggaaactttctctgggacctggcaggtggacagacacttaattggaaatt 1922
 QY 181 aacaccttggggaacacattgtgacacggagggtgcggcaagcctataaagcctac 240
 Db 1923 aatacactgggagaacacattgtctgataatggaggtctgttcaagcatacagagcctat 1982
 QY 241 ctcaagtggatggcagagggtggcaggaccagcagctgcccggcctggatctcaccatc 300
 Db 1983 cagaattattataaaagaatggcgaagaaaaattacttctctggacttgacctaaatcac 2042
 QY 301 gacagctcttcttcaactatgc 326
 Db 2043 aaacaactattttcttgaaacttgc 2068

Search completed: October 6, 2002, 02:37:29
 Job time: 4171 sec

1	128.4	39.3	570	3	US-08-646-273-13	Sequence 13, Appl
2	128.4	39.3	2129	3	US-08-646-273-22	Sequence 22, Appl
3	128.4	39.3	2314	3	US-08-646-273-29	Sequence 29, Appl
4	127.4	39.0	2889	1	US-08-289-113-1	Sequence 1, Appl
5	126.2	38.6	1703	3	US-08-646-273-18	Sequence 18, Appl
6	120.4	36.8	2533	3	US-08-646-273-24	Sequence 24, Appl
7	120.4	36.8	2720	3	US-08-646-273-35	Sequence 35, Appl
8	107.32	31.7	3291	1	US-08-574-763-1	Sequence 1, Appl
9	103.8	31.7	2560	4	US-09-305-640-1	Sequence 17, Appl
10	37.4	11.4	146	1	US-08-337-268A-54	Sequence 54, Appl
11	37.4	11.4	146	1	US-08-484-570A-54	Sequence 54, Appl
12	37.4	11.4	2576	1	US-08-471-033-35	Sequence 35, Appl
13	37.4	11.4	2576	2	US-08-471-044-35	Sequence 35, Appl
14	37.4	11.4	2576	2	US-08-463-483A-35	Sequence 35, Appl
15	37.4	11.4	2576	2	US-08-471-046A-35	Sequence 35, Appl
16	37.4	11.4	2576	2	US-08-470-566B-35	Sequence 35, Appl
17	37.4	11.4	2576	2	US-08-469-334-35	Sequence 35, Appl
18	37.4	11.4	2576	3	US-09-300-529-35	Sequence 35, Appl
19	37.4	11.4	2655	1	US-08-471-033-17	Sequence 17, Appl
20	37.4	11.4	2655	1	US-08-471-033-26	Sequence 26, Appl
21	37.4	11.4	2655	2	US-08-471-044-17	Sequence 17, Appl
22	37.4	11.4	2655	2	US-08-471-044-26	Sequence 26, Appl
23	37.4	11.4	2655	2	US-08-463-483A-17	Sequence 17, Appl
24	37.4	11.4	2655	2	US-08-463-483A-26	Sequence 26, Appl
25	37.4	11.4	2655	2	US-08-471-046A-17	Sequence 17, Appl
26	37.4	11.4	2655	2	US-08-471-046A-26	Sequence 26, Appl
27	37.4	11.4	2655	2	US-08-470-566B-17	Sequence 17, Appl


```

; TELEFAX: (516) 822-3582
;
; INFORMATION FOR SEQ ID NO: 5
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 146 nucleotides
;
; TYPE: nucleic acid
;
; STRANDEDNESS: single
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: genomic DNA
;
; HYBOTYPING: no
;
; US-08-337-258A-54

```

```
Query Match      11.4%; Score 37.4; DB 1; Length 146;
Best Local Similarity 62.1%; Pred. No. 0.033;
Matches 59; Conservative 0; Mismatches 36; Indels
```

Qy 233 aggcctactcaagtggatgagcagagggttgcaaggaccagcagctgcccgccctggatc...29
||||| |||| | || | | ||| ||||| ||||| |
Db 24 AGGCATACAGCAAGAGGCTGTATTCCGCCACCATGGGAGACTGCTCGCCCCAGCTGGACC...83
||||| |||| | || | | ||| ||||| ||||| |

Qy 293 tcacccatgagcagctcttcttcatacaactatgcc 327
 ||| || ||||| ||||| ||||| |||||
 Db 84 TCAGCCCCCAGCAGATCTTCTTTTCGAGCTATGCC 118

RESULT 11

US-08-484-570A-54
; Sequence 54, Application US/08484570A
; Patent No. 5804379
; GENERAL INFORMATION:

APPLICANT: Lee, Soohae
APPLICANT: Redman, Colvin L.
TITLE OF INVENTION: Diagnostic Method and Kit for
TITLE OF INVENTION: Determining Kell Blood Group
TITLE OF INVENTION: Genotype
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:

CORRESPONDENCE ADDRESS:
ADDRESS: Ronald J. Baron, Esq.
ADDRESS: Hoffmann & Baron
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: USA
ZIP: 11753

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484.570A
FILING DATE: 07-JUN-1995

```

; FILING DATE: 07 JUN 1993
 ; CLASSIFICATION: 435
 ;
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/337,268
 ; FILING DATE: 11-OCT-1994
 ; CLASSIFICATION: 435
 ;

ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 454-3 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582

```

/ LIBRARY: (J01) 022-3582
/ INFORMATION FOR SEQ ID NO: 54:
/
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 146 nucleotides
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: genomic DNA
/ HYPOTHEetical: no
/ US-08-484-570A-54

```

US-08-484-570A-54

	Query Watch	11.4%	Score 37.4	DB 1	Length 146
	Best Local Similarity	62.1%	Pred No. 0.033		
	Matches 59	Conservative 0	Mismatches 36	Indels 0	Gaps 0
QY	233	aggcctacctcaagtggatggcagaggggtggcaaggaccagcagctgccggcctgatac	292		
DB	24	AGCATACACAGAGAGGCTTTACGGCACCATGGGGAGACTGCTCCGCCAGCTGGACC	83		
QY	293	tcaccatgatgacgagctctcttcatcaactatgcc	327		
DB	84	TCAGCCCCCAGCAGATCTTCTTCGAAGCTATGCC	118		

RESULT · 12

RESOL 12
US-08-471-033-35
; Sequence 35, Application US/08471033
: Patent No. 5770696

GENERAL INFORMATION:

APPLICANT: Warren, Gregory W
APPLICANT: Koziel, Michael G
APPLICANT: Mullins, Martha A
APPLICANT: Nye, Gordon J
APPLICANT: Carl, Brian
APPLICANT: Desai, Nallini M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Estruch, Juan J
TITLE OF INVENTION: NO. 5770896el P
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:

ADDRESS: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532

ZIP: 10532
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471-033

ATTENTION NUMBER: 00/00/472/0000
FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA: US 08/314,594
APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/037,057
FILING DATE: 25-MAR-1993
ATTORNEY/AGENT INFORMATION:

ATTORNEY/AGENCY INFORMATION.
NAME: Pace, Gary M.

NAME: RACE, GUY M.
REGISTRATION NUMBER: P-40,403
REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV7 - SQLV3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582

TELEFAX: 919-541-8689

```

INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
    LENGTH: 2576 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: 9..2564
; OTHER INFORMATION: /note= "Maize optimized sequence
; OTHER INFORMATION: encoding VIPIA(a) with the Bacillus secretion signal removed
; OTHER INFORMATION: contained in PCIB5526"
US-08-471-033-35

Query Match 11.4%; Score 37.4; DB 1; Length 2576;
Best Local Similarity 46.1%; Pred. No. 0.094;
Matches 125; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

Qy 36 ccggaacttcgacagaatgcaacatgatgatgggtggtggaatattctcaccaccagca 95
Db 491 CCCCGAGTTCAACAAGAGGAGGAGGAGGAGTTCTGCGCAAGCCGACGCAAGATCAACCT 550
Qy 96 cttccggagcagtcagatgatgatctaccagtcacgcaactctctgggacctggc 155
Db 551 GTTACCCAGCAGATGAAGCGGAGATCGACGAGACCCGACACCGGCGGACAGCAT 610
Qy 156 agacgaacagacgtgaacggattcaacaccccttgggaaacattgtgacacggagg 215
Db 611 CCCCGACCTGTGGGAGGAGAACGGCTACACCATCCAGAACCCGCTGCGTGAAGTGGGA 670
Qy 216 ggtcggaagcctaagcctaccctcaagtggatggcagaggggtggaagaccagca 275
Db 671 CGACACCTGGCTAGCAGGGCTACACCAAGTTCGTGAGCAACCCCTCGGAGAGCCACAC 730
Qy 276 gctgcccggcctggatctcaccatgagcag 306
Db 731 CGTGGCGGACCCCTACACCGACTACGAGAG 761

RESULT 13
US-08-471-044-35
; Sequence 35, Application US/08471044
; Patent No. 5840868
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5840868el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,044
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,483
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
```

```

; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV6 - SOLv3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 9..2564
; OTHER INFORMATION: /note= "Maize optimized sequence
; OTHER INFORMATION: encoding VIPIA(a) with the Bacillus secretion signal remov
; OTHER INFORMATION: contained in PCIB5526"
US-08-471-044-35

Query Match 11.4%; Score 37.4; DB 2; Length 2576;
Best Local Similarity 46.1%; Pred. No. 0.094;
Matches 125; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

Qy 36 ccggaacttcgacagaatgcaacatgatgatgggtggtggaatattctcaccaccagca 95
Db 491 CCCCGAGTTCAACAAGAGGAGGAGGAGTTCTGCGCAAGCCGACGCAAGATCAACCT 550
Qy 96 cttccgggagcagtcagatgatgatctaccagtcacgcaactctctgggacctggc 155
Db 551 GTTACCCAGCAGATGAAGCGGAGATCGACGAGACCCGACACCGGCGGACAGCAT 610
Qy 156 agacgaacagacgtgaacggattcaacaccccttgggaaacattgtgacacggagg 215
Db 611 CCCCGACCTGTGGGAGGAGAACGGCTACACCATCCAGAACCCGCTGCGTGAAGTGGGA 670
Qy 216 ggtcgcggaagcctataagcctaccctcaagtggatggcagaggggtggaagaccagca 275
Db 671 CGACACCTGGCTAGCAGGGCTACACCAAGTTCGTGAGCAACCCCTCGGAGAGCCACAC 730
Qy 276 gctgcccggcctggatctcaccatgagcag 306
Db 731 CGTGGCGGACCCCTACACCGACTACGAGAG 761

RESULT 14
US-08-463-483A-35
; Sequence 35, Application US/08463483A
; Patent No. 5849870
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5849870el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
```


CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463.483A
FILING DATE: 09-SEP-1994
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/037,057
FILING DATE: 25-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Spruill, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: CGC 1695/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8615
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 2576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 9..2564
OTHER INFORMATION: /note= "Maize optimized sequence
encoding VIP1(a) with the Bacillus secretion signal removed
contained in pCIB5526"
US-08-463-483A-35

Query Match 11.4%; Score 37.4; DB 2; Length 2576;
Best Local Similarity 46.1%; Pred. No. 0.094;
Matches 125; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
QY 36 ccggaacttcacagaatgcaacatgatgagtcgggagtaacttctccaccagca 95
Db 491 CCCGAGTTCACAAAGAGAGAGAGAGAGTCTGCGCAAGCCGACGATCAACCT 550
QY 96 ctccggagcagtcagatgatctaccagtcagcgaactctctggagcctggc 155
Db 551 GTTACCCAGCAGATGAGCGGAGATCGACGAGACACCGACCGCGGACGAT 610
QY 156 agacgaacagacgtgaacggtattcaacaccccttgggaaacattgtgacaacggag 215
Db 611 CCCCGACCTGTGGGAGGAGAGACGGGTACACCATCAGCAACCGCATCGCGTGAAGTGGA 670
QY 216 ggtgcggaacacctataaggctactcaagtggatggcagaggggtgcaagaccagca 275
Db 671 CGACAGCTGGCTAGCAAGGGCTACCAAGGCTACCAAGTTCGTGAGCAACCCCTGGAGAGCCAC 730
QY 276 gctgcggcgcctggatctcaccatgagcag 306
Db 731 CGTGGCGCCACCCCTACACCGACTACGAGAAG 761

RESULT 15

US-08-471-046A-35
Sequence 35, Application US/08471046A
Patent No. 5866326
GENERAL INFORMATION:
APPLICANT: Warren, Gregory W
APPLICANT: Koziel, Michael G
APPLICANT: Mullins, Martha A
APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian
APPLICANT: Desai, Nallini M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Estruch, Juan J
TITLE OF INVENTION: Method For Isolating Vegetative Insecticidal
Protein Genes
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5866326artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471.046A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/463.483
FILING DATE: 05-JUN-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/037,057
FILING DATE: 25-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC1695/CIP3/DIV8 - SOLV4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 2576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 9..2564
OTHER INFORMATION: /note= "Maize optimized sequence
encoding VIP1(a) with the Bacillus secretion signal removed
contained in pCIB5526"
US-08-471-046A-35

Query Match 11.4%; Score 37.4; DB 2; Length 2576;
Best Local Similarity 46.1%; Pred. No. 0.094;
Matches 125; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

```

QY 36 ccggaacttcgacaagaatggcaacatgatggattgggagtaactttctccaccagca 95
Db 491 CCCCGAGTTCAACAAGAAGGAGAGCCAGGAGTTCTTGCCCAAGCCAGCAAGATCAACCT 550
QY 96 cttccgggagcagtcagtgatgatctaccagtaacggaactactccttgggacctggc 155
Db 551 GTTACCCAGCAGATGAAGCGCGAGATCGACGAGGACACCGACCCGCGACAGCAT 610
QY 156 agacgaacagaacgtgaacggattcaacacccttgggaaacattgctgacaacggagg 215
Db 611 CCCCGACCTTGGGAGAGAGAACCGGTACACCATCCAGAACCGCATCCCCGTGAAGTGGGA 670
QY 216 ggtcgcgcaagcctataaaggcctacctaagtgatggcagaggggtggcaaggaccagca 275
Db 671 CGACAGCCTGGCTAGCAAGGGCTACACCAAGTTCGTGAGCAACCCCTGGAGAGCCACAC 730
QY 276 gctgccggcctggtatctaccatgagcag 306
Db 731 CGTGGCGACCCCTACACCGACTACGAGAAG 761

```

Search completed: October 6, 2002, 02:01:38
Job time: 4030 sec

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2002, 00:27:23 ; Search time 1714.12 Seconds
(without alignments)
2574.793 Million cell updates/sec

Title: US-09-647-780A-3
Perfect score: 327
Sequence: 1 gggcagagatcacgcacgg.....tcttttcacatcatgccc 327

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: EST:

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estm:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_gss:*
- 13: em_gss_hum:*
- 14: em_gss_inv:*
- 15: em_gss_pln:*
- 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	165.6	50.6	667	9 AJ396658	AJ396658
2	159.8	48.9	531	10 BI828288	BI828288
3	159.8	48.9	586	10 BI907511	BI907511
4	158	48.3	1144	10 BM459254	BM459254
5	147.8	45.2	610	9 AW465461	AW465461
6	142.6	43.6	926	10 BF968576	BF968576
7	142	43.4	705	9 AI325325	AI325325
8	139.2	42.6	550	10 BE245519	BE245519
9	138.2	42.3	549	10 BE245100	BE245100
10	138.2	42.3	899	9 AW940129	AW940129
11	132.6	40.6	454	9 AI671752	AI671752
12	132.6	40.6	540	10 BF476098	BF476098
13	132	40.4	391	10 BF590024	BF590024
14	131	40.1	476	9 AI702551	AI702551
15	128.2	39.2	443	9 AI566727	AI566727
16	126.4	38.7	563	9 AW086484	AW086484
17	126	38.5	458	9 AI913156	AI913156

18	123.6	37.8	501	10 W89492	W89492
19	123.6	37.8	536	10 BI692693	BI692693
20	120.6	36.9	466	10 WI0214	WI0214
21	120.4	36.8	968	10 BI869244	BI869244
22	120	36.7	517	10 BI326165	BI326165
23	119.8	36.6	660	10 BM491749	BM491749
24	118.8	36.3	423	9 AA895479	AA895479
25	117.4	35.9	479	9 AA450725	AA450725
26	116.8	35.7	485	10 W78610	W78610
27	113.4	34.7	608	10 BI360370	BI360370
28	112.2	34.3	693	10 BE535389	BE535389
29	108	33.0	704	10 BG153322	BG153322
30	107.6	32.9	600	10 BG803768	BG803768
31	107.6	32.9	600	10 BI986454	BI986454
32	107.4	32.8	548	9 AW654249	AW654249
33	107.4	32.8	761	10 BE871194	BE871194
34	106.8	32.7	678	10 BI505516	BI505516
35	106.4	32.5	463	10 W48775	W48775
36	106.2	32.5	249	9 AI874916	AI874916
37	106	32.4	626	10 BJ093321	BJ093321
38	106	32.4	651	10 BJ096397	BJ096397
39	105.6	32.3	592	10 BF044240	BF044240
40	104.2	31.9	530	10 BF562978	BF562978
41	103.8	31.7	889	10 BF663245	BF663245
42	101.2	30.9	665	9 AV329959	AV329959
43	99.6	30.5	905	10 BF124139	BF124139
44	99.2	30.3	357	9 AA026997	AA026997
45	98.6	30.2	822	10 BG283366	BG283366

ALIGNMENTS

RESULT 1
AJ396658 667 bp mRNA linear EST 25-JAN-2001
LOCUS
DEFINITION AJ396658 dkfz426 Gallus gallus cDNA clone 27e4rl, mRNA sequence.
ACCESSION AJ396658
VERSION AJ396658.1 GI:7128817
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 667)
AUTHORS Adrakhmanov, I., Lodygin, D., Geroth, P., Arakawa, H., Law, A., Plachy, J., Korn, B. and Buerstedde, J.M.
TITLE A large database of chicken bursal ESTs as a resource for the analysis of vertebrate gene function
JOURNAL Genome Res. 10 (12), 2062-2069 (2000)
MEDLINE 20568495
COMMENT Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
Location/Qualifiers
1. 667
/organism="Gallus gallus"
/strain="CB"
/db_xref="taxon:9031"
/clone="27e4rl"
/clone_lib="dkfz426"
/tissue_type="Bursa of Fabricius"
BASE COUNT 219 a 121 c 165 g 160 t 2 others
ORIGIN

Query Match 50.6%, Score 165.6; DB 9; Length 667;
Best Local Similarity 69.0%; Pred. No. 3.7e-33;
Matches 225; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

```

QY 1 ggcacgagatcacgcgctttgatgacaaatggccgggaacttcgacagaatggcaac 60
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 320 GGTCAAGAAATACACACGGTTTGTATGACACGGCAGAAATTTCAATGAGATGGAGAC 379
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 atgatggattggaggaactctctccacccagcaacttcgaggagcagtgatcagatg 120
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 380 CTGATGACTGGTGGAGTGAAGAAATCTGCACGTAACTTTAAGGACTTTATCGCACTGCATG 439
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 ataccagtcacggaactactctctggaacttcgacagcaagaacagtggaacgattc 180
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 440 GTGTACCAAGTATGAAACTTCTCATGGGACCTAGCAGGTGGACAGCAGCTGAGTGGAAATC 499
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 aacaccttgggggaaacattctgacaaaggagggtgcggaagcctataaagccctac 240
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 500 AACACACTAGGAGAAACATTTCTGATAACGGGGTGTATAGCAAGCATATAAGGCCAT 559
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 ctcaagtggatgcagagggtgccaagaccagcagctgcccggcctggatctcaccat 300
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 560 GAGAACTTTCTGAAAAAGAAATGGGAAAGAAAACTTCTCCCTGNTCTGGACATGAACCAT 619
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 gacgactctcttcacactatgc 326
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 620 CACGACTGNTTTTCTCAACTTGC 645
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
BI828288
LOCUS BI828288 531 bp mRNA linear EST 04-OCT-2001
DEFINITION 603078182F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169765 5',
mRNA sequence.
ACCESSION BI828288
VERSION BI828288.1 GI:15939838
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 531)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabp@remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1422 row: d column: 22
High quality sequence stop: 512.

FEATURES
source
1..531
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5169765"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

BASE COUNT 115 a 163 c 173 g 80 t
ORIGIN

```

```

Query Match 48.9%; Score 159.8; DB 10; Length 531;
Best Local Similarity 98.8%; Pred. No. 1.1e-31;
Matches 161; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 165 gaacgtgaacggattcaacaccttgggggaaacattctgacaaacgaggggtgcgca 224
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 110 GCAGGTGAACGGATTCAACACCCCTTGGGGAAAAACATTGCTCACAACGAGGGTGCAGCA 169
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 225 agcctataaaggcctacactcaagtggatggcagaggtggcagaccagcagctgcccgg 284
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 170 AGCCTATAGCCCTACCTCAAGTGGTGGCAGAGGGTGGCAAGGACACGACGAGCTGCCGG 229
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 285 cctggatctcaccatcagcagctctcttcacactatgcc 327
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 230 CCTGGATCTCACCCTCAGCAGCTCTTCTTTCATCAACTATGCC 272
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
BI907511
LOCUS BI907511 586 bp mRNA linear EST 16-OCT-2001
DEFINITION 603065633F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5214609 5',
mRNA sequence.
ACCESSION BI907511
VERSION BI907511.1 GI:16170345
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 586)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabp@remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1539 row: a column: 10
High quality sequence stop: 551.

FEATURES
source
1..586
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5214609"
/clone_lib="NIH_MGC_118"
/tissue_type="leukocyte"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."

BASE COUNT 132 a 180 c 184 g 90 t
ORIGIN

Query Match 48.9%; Score 159.8; DB 10; Length 586;
Best Local Similarity 98.8%; Pred. No. 1.2e-31;
Matches 161; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 165 gaacgtgaacggattcaacaccttgggggaaacattctgacaaacgaggggtgcgca 224
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 156 GCAGGTGAACGGATTCAACACCCCTTGGGGAAAAACATTGCTCACAACGAGGGTGCAGCA 215
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Query Match	48.3%;	Score 158;	DB 10;	Length 1144;
Best Local Similarity	67.8%;	Pred. No. 4.5e-31;		
Matches 221;	Conservative 0;	Mismatches 105;	Indels 0;	Gaps 0;
1	g g g c a c g a g a t c a c g a c g c t t t g a t g a c a a t g g c c g a a c t t c g a c a g a a t g g c a a c	60		
90	g g a c a c g a a a t c a c c c a t g c t t c g a t g a c a a t g g c a g a a a c t t a c a a a g a t g g a g a c	149		
61	a t g a t g g a t t g g t g g a g t a a c t t c t c a c c a c g a c a c t t c c g g a g c a g t c a g a g c a t g	120		
150	c t c g t t g a c t g g t g g a c t c a a c a g t c t g c a a g t a a c t t t a a g g a g c a a t c c c a g t g c a t g	209		
121	a t c t a c c a g t a c g c g a a c t a c t c c t g g a c c t g g c a g a c g a a c g a a c g t g a a c g g a t t c	180		
210	g t g t a t c a g t a t g g a a a c t t t t c c t g g g a c c t g c a g t g g a c a g c a c c t a a t t g g a a t t	269		
181	a a c a c c c t t g g g g a a a c a a t t g t g a c a a c g g g g g t g c g g c a a g c c t a a a g g c c t a c	340		
270	a a t a c a t c g g g a a a a c a t t g c t g a t a t g g a g g t t t t g c t c a a g c a a t a c a g a g c c t a t	329		

FEATURES

```
1: .610
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="np23019B10E11"
/clone.lib="Soares normal"
/sex="female"
/lab_host="DH10B"
/notes="Organ: placenta;
Site: 2; Motif: The cDNA
Soares laboratory and i
as described by Bonaldo
M.B. (1996), Genome Res
194 a 124 c 139 q 1
```

	BASE COUNT	ORIGIN
1	0.000	0.000
2	0.000	0.000
3	0.000	0.000
4	0.000	0.000
5	0.000	0.000
6	0.000	0.000
7	0.000	0.000
8	0.000	0.000
9	0.000	0.000
10	0.000	0.000
11	0.000	0.000
12	0.000	0.000
13	0.000	0.000
14	0.000	0.000
15	0.000	0.000
16	0.000	0.000
17	0.000	0.000
18	0.000	0.000
19	0.000	0.000
20	0.000	0.000
21	0.000	0.000
22	0.000	0.000
23	0.000	0.000
24	0.000	0.000
25	0.000	0.000
26	0.000	0.000
27	0.000	0.000
28	0.000	0.000
29	0.000	0.000
30	0.000	0.000
31	0.000	0.000
32	0.000	0.000
33	0.000	0.000
34	0.000	0.000
35	0.000	0.000
36	0.000	0.000
37	0.000	0.000
38	0.000	0.000
39	0.000	0.000
40	0.000	0.000
41	0.000	0.000
42	0.000	0.000
43	0.000	0.000
44	0.000	0.000
45	0.000	0.000
46	0.000	0.000
47	0.000	0.000
48	0.000	0.000
49	0.000	0.000
50	0.000	0.000
51	0.000	0.000
52	0.000	0.000
53	0.000	0.000
54	0.000	0.000
55	0.000	0.000
56	0.000	0.000
57	0.000	0.000
58	0.000	0.000
59	0.000	0.000
60	0.000	0.000
61	0.000	0.000
62	0.000	0.000
63	0.000	0.000
64	0.000	0.000
65	0.000	0.000
66	0.000	0.000
67	0.000	0.000
68	0.000	0.000
69	0.000	0.000
70	0.000	0.000
71	0.000	0.000
72	0.000	0.000
73	0.000	0.000
74	0.000	0.000
75	0.000	0.000
76	0.000	0.000
77	0.000	0.000
78	0.000	0.000
79	0.000	0.000
80	0.000	0.000
81	0.000	0.000
82	0.000	0.000
83	0.000	0.000
84	0.000	0.000
85	0.000	0.000
86	0.000	0.000
87	0.000	0.000
88	0.000	0.000
89	0.000	0.000
90	0.000	0.000
91	0.000	0.000
92	0.000	0.000
93	0.000	0.000
94	0.000	0.000
95	0.000	0.000
96	0.000	0.000
97	0.000	0.000
98	0.000	0.000
99	0.000	0.000
100	0.000	0.000

```
Query Match      45.2%; Score 147.8; DB 9; Length 610;
Best Local Similarity 65.7%; Pred. No. 1.7e-28;
Matches 215; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
```

```

QY 61 atgattggtggaagtaactctccaccacagcacttcggggagcagtgatgcatg 120
DB 278 CTTGTTGACTGGTGGAGTCAACAGTCTGCAATAATTTTAAAGACCTATCCAGTGCATG 337
QY 121 atctaccagtacggaactactctcgggacgtggcagacgaacagaacggtggaacgattc 180
DB 338 GTGTACAGTATGGGAACCTTCTCTGGGACCTAGCAAAATGGACAGCATCTCAATGGGAATT 397
QY 181 aacacccttggggaacattctgacaacgaggggtgcggaagcctataaagcctac 240
DB 398 AATACACTGGGAGAAACATCTGTAATGCTGTATGGCCAAAGCATACAGAGCCTAT 457
QY 241 ctcaagtggatggcagaggggtggcaagaccagcagctgccggcctggatctcaaccat 300
DB 458 CAAACATATGTTAAAGCATGGTGAAGAAAAATTTACTTCTGGACITGACCTAAATCAC 517
QY 301 gacagctcttcttcatacaactatgcc 327
DB 518 AGACAACCTGTTCTCTCGAATTTTGC 544

RESULT 6
BF968576 926 bp mRNA linear EST 22-JAN-2001
LOCUS 602271182F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4359141 5'
DEFINITION mRNA sequence.
ACCESSION BF968576
VERSION BF968576.1 GI:12335791
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL 1 (bases 1 to 926)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-femail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM9998 row: d column: 22
High quality sequence stop: 766.
FEATURES
source
1..926
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4359141"
/tissue_type="adrenal cortex carcinoma, cell line"
/note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site: 1.
Noti: Site 2: SalI; Cloned unidirectionally; oligo-dr
primed. Average insert size 1.229 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 286 a 192 c 221 g 227 t
ORIGIN
Query Match 43.6%; Score 142.6; DB 10; Length 926;
Best Local Similarity 67.3%; Pred. No. 4.5e-27;
Matches 216; Conservative 0; Mismatches 104; Indels 1; Gaps 1;

QY 6 cgagatcacgcacgctttgatgacaatggccggaacttcgacaagaatggcaacatgat 65
DB 1 CGAATCACCCATGCTTCGATGACAAATGGGAGAAACCTTTAACAAAGATGGAGACCTCGT 60

```

```

QY 66 ggattggtggaagtaactctccaccacagcacttcggggagcagtgatgcatg 125
DB 61 TGACTGTGGGACTCAACAGTCTGCAAGTAACCTTTAAGGAGCAATCCCACTGATGGTCTGA 120
QY 126 ccagtacgggaactactctcctgggacgtggcagacgaacagaacggtggaacgattcaaac 185
DB 121 TCAGTATGGAAACCTTTCTCTGGGACCTGGCAGGTGGACAGCACCTTAATGGAATTAATAC 180
QY 186 ccttggggaacattctgacaacgaggggtgcggaagcctataaagcctaccta 245
DB 181 ACTGGGAGAAACA-TGCTGATATGGAGGCTCTGGTCAAGCATACAGAGCCTATCAGAA 239
QY 246 gtgatggcagaggtggcagagaccagcagctgccggcctggatctcaaccatgagca 305
DB 240 TTATATTAAGAAAGTGGCGAAGAAAAATTTACTTCTGGACTTGACCTAAATCACAACA 299
QY 306 gctcttcttcatacaactatgc 326
DB 300 ACTATTTTCTTGAACCTTGC 320

RESULT 7
AI325325 705 bp mRNA linear EST 23-DEC-1998
LOCUS ma44c03.y1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
DEFINITION IMAGE:313540 5' similar to gb:J03779 NEPRILYSIN (HUMAN);, mRNA
sequence.
ACCESSION AI325325
VERSION AI325325.1 GI:4059754
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 705)
COMMENT Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:204156
This read is a RESEQUENCE of a previously sequenced mouse clone.
This read has been verified (found to hit its original self in the
correct orientation)
Seq primer: -40RP from Gibco
High quality sequence stop: 459.
FEATURES
source
1..705
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:313540"
/dev_stage="19.5 dpc total fetus"
/note="Vector: DH10B (ampicillin resistant)"
/lab_host="DH10B (Pharmacia) with a modified
polylinker; Site: 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGGAGCGCCGCGCATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of

```


full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Seq primer: M13 primer.

FEATURES

source
Location/Qualifiers
1..549
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="TCBAP2535"
/clone_lib="Pediatric pre-B cell acute lymphoblastic leukemia Baylor-HGSC project-TCBA"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="pre-B cell"
/dev_stage="pediatric 2 years"
/lab_host="DH10B"
/note="Vector: lambda PSB; Site.1: BamHI; Site.2: EcoRI; First strand cDNA was primed with an anchored XhoI-oligo(dT) primer [5'GGAGACTCGAGCGCCGACGAGAG(T)VN 3'; V-A,C,G; N-A,C,G,T] and then dG tailed. Second strand was primed with a BamHI-dC primer [5'AGAGACTCGGATCCGCGCCGCAATAAATAAT(C) 3']. Double-stranded cDNA was then digested with BamHI and XhoI and directionally cloned into the BamHI and SalI sites of lambda PSB-vector. Library went through one round of normalization. Library was constructed by Wei Yu at RIKEN of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T, Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M, Schneider C, Hayashizaki Y, High efficiency selection of full-length cDNA by improved biotinylated cap trapper., DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT 183 a 112 c 124 g 129 t 1 others
ORIGIN

Query Match 42.3%; Score 138.2; DB 10; Length 549;
Best Local Similarity 67.4%; Pred. No. 5.2e-26;
Matches 194; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 1 ggcacagatcagcagcgtttgatgacaaatggccggaacttcgacagaatggcaac 60
Db 249 GGACAGAAATACCCATGCTTCGATGACAAATGGCAGAACTTTAAAGATGGAGAC 308
QY 61 atgatggattggagtaactttctccaccagcacttcctgggagcagtcagagtgcag 120
Db 309 CTCGTTGACGTGGGACTCAACAGTCGTGCAAGTAACCTTAAGGAGCAATCCAGTCGATG 368
QY 121 atctaccagtcagcgaactactctctgggaactggcagacgaagaacgtaacggttc 180
Db 369 GTGTATCAGTATGGAATCTTTCTCGGACCTGGCAGGTGGACAGCACCTTAATGGAATT 428
QY 181 aacaccttgggaaacattgtgacacgaggggtgcggaagcctataagcctac 240
Db 429 AATACACTGGGAGAAACATCTGCTGATANTGGAGCTTTGGTCAAGCATACAGAGCCTAT 488
QY 241 ctcaagtggatggcagaggtggcgaagaccagcagctgccggcctg 288
Db 489 CAGATTATATANAAGAATGGGAGAAATAATTAATCTCTGGACTTG 536

RESULT 10
AW940129/c
LOCUS
DEFINITION
AW940129 899 bp mRNA linear EST 23-APR-2001
melanogaster cDNA clone GH03315 3, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
fruit fly.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 899)
REFERENCE
AUTHORS
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,

TITLE

BDGP/HMI Drosophila EST Project
Unpublished (2001)
Other_ESTs: GH03315.5prime
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798

Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or more T residues at the beginning of the sequence, this clone was polyadenylated. The resulting Poly-T sequence has been removed. hit genomic AE003436: Drosophila melanogaster genomic scaffold 142000013386054 section 20 of 35, complete sequence.: 02/14/2001
Plate: GH.33 row: B column: 3
High quality sequence stop: 819.
Location/Qualifiers
1..899

FEATURES

source
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="GH03315"
/clone_lib="GH Drosophila melanogaster head pOT2"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"
/note="Organ: head; Vector: pOT2; Site.1: EcoRI; Site.2: XhoI; Sized fractionated cDNAs were directly ligated into pOT2. Plasmid cDNA library."
BASE COUNT 213 a 223 c 240 g 221 t 2 others
ORIGIN

Query Match 42.3%; Score 138.2; DB 9; Length 899;
Best Local Similarity 66.1%; Pred. No. 6.2e-26;
Matches 216; Conservative 0; Mismatches 108; Indels 3; Gaps 1;
QY 1 ggcacagatcagcagcgtttgatgacaaatggccggaacttcgacagaatggcaac 60
Db 830 GGCCACGAGATTACCCACGGATTGACGACAAAGGCGAGGCAATTCGACAAAGGAGGCAAC 771
QY 61 atgatggattggagtaactttctccaccagcacttcctgggagcagtcagagtgcag 120
Db 770 ATGATGAGTGGTGGAAACAATGCCACCATCGAGGCTTTTCGGAACGACGAGTGCCTC 711
QY 121 atctaccagtcagcgaactactctctgggaacttcctccaccagcacttcctgggagcagtcagagtgcag 180
Db 710 ATCGATCAGTACTCGCGCTACAAGATTAAACGAGGTGGAC---ATGTTTCATGACGCGGCGG 654
QY 181 aacaccttgggaaacattgtgacacgaggggtgcggaagcctataagcctac 240
Db 653 ATGACGAGGCGGAGACATCGCGACAAACGCGGCTCTAAACAGGCTTTTCAGGCGCTAC 594
QY 241 ctcaagtggatggcagaggtggcgaagaccagcagctgccggcctggtatctccaccat 300
Db 593 AAGAAATGGGAGACCTTACATGGGCGGAGCAGCAGCTGCCCGGCTTGAACATGATCCAC 534
QY 301 gacagctctcttcacatcaactatgcc 327
Db 533 GATCAGCTGTTCTTTCTCACTACGCC 507

RESULT

AW1671752/c
LOCUS
DEFINITION
AW1671752 454 bp mRNA linear EST 16-DEC-1999
similar to gb:J03779 NEPRILYSIN (HUMAN);, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 454)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP);
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone Distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/bbrp/image/image.html
 Insert Length: 622, Std Error: 0.00
 Seq primer: -400P from Gibco.
 Location/Qualifiers
 1..454

FEATURES
 source
 1..454
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2297228"
 /clone_lib="NCI-CGAP_Kid11"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: p7T3D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
 Plasmid DNA from the normalized library NCI-CGAP_Kid3 was
 prepared, and ss circles were made in vitro. Following HAP
 purification, this DNA was used as tracer in a subtractive
 hybridization reaction. The driver was PCR-amplified cDNAs
 from a pool of 5,000 clones made from the same library
 (cloneIDs 1322376-1323911, 1456007-1456775, and
 1500552-1502855). Subtraction by Bento Soares and M.
 Fatima Bonaldo."
 BASE COUNT 106 a 111 c 95 g 142 t
 ORIGIN

Query Match 40.6%; Score 132.6; DB 9; Length 454;
 Best Local Similarity 70.5%; Pred. No. 1.4e-24;
 Matches 177; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
 QY 1 ggcacagatcacgcagcgctttgatgacaatggccggaacttcgacagaatggcaac 60
 Db 252 GGACACGAATATCCCATGGCTTCGATGACAATGGCAGAACTTTAACAAAGATGGAGAC 193
 QY 61 atgatgattgttgagtaacttctccaccagcacttcggagcagtcagagtgcattg 120
 Db 192 CTCGTGACTGGTGGACTCAACAGCTCTGCAAGTAACCTTTAAGGAGCAATCCAGTGCATG 133
 QY 121 atctaccagtcagcgaactactcctgggacctggcagcagcaacgagcgtgaacgattc 180
 Db 132 GTGTATCAGTATGAAACTTTCTCTGGACCTGGCAGCTGGACGACCTTAATGGAATT 73
 QY 181 aacacccctggggaacattgtgacaacggaggggtgcggcaagcctataagcctac 240
 Db 72 AATACACTGGGAGAAACATTTGCTGATAATGGAGGCTTGTGTCAGCATACAGAGCCTAT 13
 QY 241 ctcaagtggat 251
 Db 12 CAGAAATTATAT 2

RESULT 12
 BF476098/c
 LOCUS hr77d03.x1 NCI-CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3134501 3'
 DEFINITION similar to SW:NEP_HUMAN P08473 NEPRILYSIN ;, mRNA sequence.
 ACCESSION BF476098
 VERSION BF476098.1 GI:11546925
 KEYWORDS EST.
 SOURCE human.

ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 540)
 REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP);
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone Distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL, send email to:
 info@image.llnl.gov
 Seq primer: -400P from Gibco
 High-quality sequence stop: 473.
 Location/Qualifiers
 1..540

FEATURES
 source
 1..540
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3134501"
 /clone_lib="NCI-CGAP_Kid11"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: p7T3D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
 Plasmid DNA from the normalized library NCI-CGAP_Kid3 was
 prepared, and ss circles were made in vitro. Following HAP
 purification, this DNA was used as tracer in a subtractive
 hybridization reaction. The driver was PCR-amplified cDNAs
 from a pool of 5,000 clones made from the same library
 (cloneIDs 1322376-1323911, 1456007-1456775, and
 1500552-1502855). Subtraction by Bento Soares and M.
 Fatima Bonaldo."
 BASE COUNT 129 a 124 c 107 g 180 t
 ORIGIN

Query Match 40.6%; Score 132.6; DB 10; Length 540;
 Best Local Similarity 70.5%; Pred. No. 1.5e-24;
 Matches 177; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
 QY 1 ggcacagatcacgcagcgctttgatgacaatggccggaacttcgacagaatggcaac 60
 Db 252 GGACACGAATATCCCATGGCTTCGATGACAATGGCAGAACTTTAACAAAGATGGAGAC 193
 QY 61 atgatgattgttgagtaacttctccaccagcacttcgggagcagtcagagtgcattg 120
 Db 192 CTCGTGACTGGTGGACTCAACAGCTCTGCAAGTAACCTTTAAGGAGCAATCCAGTGCATG 133
 QY 121 atctaccagtcagcgaactactcctgggacctggcagcagcaacgagcgtgaacgattc 180
 Db 132 GTGTATCAGTATGAAACTTTCTCTGGACCTGGCAGCTGGACGACCTTAATGGAATT 73
 QY 181 aacacccctggggaacattgtgacaacggaggggtgcggcaagcctataagcctac 240
 Db 72 AATACACTGGGAGAAACATTTGCTGATAATGGAGGCTTGTGTCAGCATACAGAGCCTAT 13
 QY 241 ctcaagtggat 251
 Db 12 CAGAAATTATAT 2

RESULT 13
 BF590024/c
 LOCUS naa48g08.x1 NCI-CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3259599 3'
 DEFINITION similar to SW:NEP_HUMAN P08473 NEPRILYSIN ;, mRNA sequence.
 ACCESSION BF590024
 VERSION BF590024.1 GI:11682348

Search completed: October 6, 2002, 02:00:33
Job time: 5590. sec

BASE COUNT	117 a	100 c	86 g	140 t
ORIGIN				
Query Match	39.2%;	Score 128.2;	DB 9;	Length 443;
Best Local Similarity	65.6%;	Pred. NO. 2e-23;		
Matches 187;	Conservative 0;	Mismatches 98;	Indels 0;	Gaps 0
QY	42	cttcgacaagaatggcaacatgatgattggtggagtaacttctcacccagcacttcg	101	
Db	443	CTTTAAACAAAGATGGAGACCTCGTTGACTGGTGGACTCAACAGTCTCCRAGTAACTTAA	384	
QY	102	ggagcagtcagagtgcatgatctaccagctacggcaactactctctggagcctggcgacga	161	
Db	383	GGAGCAATCCCGACTGCATGGTGATCAGTATGGAACATTTTCTGGGACCTGGCAGGTGG	324	
QY	162	acagaacgtggaacggattcaacacccttgggaaacattgtcgacaacggaggggtgcg	221	
Db	323	ACAGCACCTTAATGGAAATTAATACCTGGGAGAAACATTGCTGATTAATGGAGGTCPTGG	264	
QY	222	gaaagcctataagcgcttacctcaagtggatggcgaggggtggcgaaggaccagcagctgcc	281	
Db	263	TCAAGCATACAGACGCTATCAGAATATATTAAGAAGATGGCGAGAAAAATTAACCTCG	204	
QY	282	cggcctggattccaccocatgagcagctctcttcatacaactatgc	326	
Db	203	TGGACTTGACCTAAATCACAACAACTATTTTCTTGAACCTTGC	159	

score_list:	Strnd	Orig	zScore	pScore	Len	Documentation
gb.pat:AX145976	+	651.00	1361.97	1.7e-67	2076	AX145976 Sequence 1 from Paten
gb.pat:AX146980	+	651.00	1361.20	1.9e-67	2262	AX146980 Sequence 5 from Paten
gb.pat:AX146978	+	651.00	1360.90	1.9e-67	2340	AX146978 Sequence 3 from Paten
gb.pat:AX139743	+	651.00	1359.83	2.2e-67	2636	AX139743 Sequence 3 from Paten
gb.pat:AX139745	+	651.00	1359.74	2.2e-67	2663	AX139745 Sequence 5 from Paten
gb.pat:AX139748	+	651.00	1359.74	2.2e-67	2663	AX139748 Sequence 5 from Paten
gb.pat:AX033274	+	651.00	1359.70	2.3e-67	2676	AX033274 Sequence 14 from Paten
gb.pat:AX139741	+	651.00	1359.57	2.3e-67	2714	AX139741 Sequence 1 from Paten
gb.pat:AXF336981	+	651.00	1359.34	2.4e-67	2784	AXF336981 Homo sapiens neprilys
gb.pat:AX139864	+	640.50	1339.00	3.2e-66	2232	AX139864 Sequence 28 from Paten
gb.pat:AX014703	+	616.00	1304.10	2.8e-64	327	AX014703 Sequence 3 from Paten
gb.ro:AF302075	+	581.00	1211.20	4.2e-59	2583	AF302075 Mus musculus neprilys
gb.ro:AF302077	+	581.00	1210.97	4.3e-59	2652	AF302076 Mus musculus neprilys
gb.ro:AF302076	+	581.00	1210.83	4.4e-59	2654	AF302077 Mus musculus neprilys
gb.ro:AF302077	+	581.00	1210.09	4.9e-59	2925	AX033272 Sequence 12 from Paten
gb.pat:AX033272	+	581.00	1210.09	4.9e-59	2925	AX033272 Sequence 12 from Paten
gb.pat:AXF163669	+	581.00	1202.10	4.9e-59	2925	AXF163669 Mus musculus neprilys
gb.pat:AX014701	+	578.00	1204.22	1.0e-58	2765	AX014701 Sequence 1 from Paten
gb.ro:AF157106	+	577.00	1202.64	1.3e-58	2801	AF157106 Mus musculus soluble
gb.ro:AF157105	+	577.00	1201.69	1.4e-58	2892	AF157105 Mus musculus soluble
gb.pat:E01557	+	471.00	974.98	6.1e-46	3369	E01557 cDNA encoding polypepti
gb.pr:HSCALLA	+	469.00	973.19	7.6e-46	2357	X00811 Human mRNA for common a
gb.om:OCNEPR	+	469.00	972.47	8.4e-46	2772	X05338 Rabbit mRNA for neutral
gb.pat:AX138810	+	469.00	971.24	9.8e-46	3181	AX138810 Sequence 2 from Paten
gb.pat:AX138995	+	469.00	971.24	9.8e-46	3181	AX138995 Sequence 2 from Paten
gb.pat:AX139414	+	469.00	971.24	9.8e-46	3181	AX139414 Sequence 2 from Paten
gb.pat:I08056	+	469.00	971.24	9.8e-46	3181	I08056 Sequence 2 from Paten
gb.pr:HSNRNEN	+	469.00	971.24	9.8e-46	3181	X07166 Human mRNA for enkephal
gb.pat:AX329921	+	469.00	966.33	1.1e-45	5508	AX329921 Sequence 430 from Pat
gb.pat:AX333461	+	469.00	966.33	1.1e-45	5508	AX333461 Sequence 3970 from Pa
gb.pat:AX336695	+	469.00	966.33	1.1e-45	5508	AX336695 Sequence 7204 from Pa
gb.pr:HUMCALLA	+	469.00	966.33	1.1e-45	5508	X03779 Human common acute lym
gb.ro:MUSCD10NEP	+	468.00	972.20	8.7e-46	2253	M81591 Murine CD10 neutral end
gb.pat:AX30431	+	465.00	974.84	6.2e-46	822	AX0431 Enkephalinase Analogue F
gb.pat:AX108057	+	465.00	967.36	3.0e-45	3243	I08057 Sequence 4 from Paten
gb.ro:RAVENKA	+	465.00	962.56	3.0e-45	3243	M1594 Rat enkephalinase (neut
gb.ov:AF077612	+	434.00	894.82	1.8e-41	3985	AF077612 Perca flavescens nepr
gb.pr:AL139246	-	397.00	783.42	2.8e-35	154736	AL139246 Human DNA sequence
gb.htg:AC014099	+	378.50	756.28	9.2e-34	39640	AC014099 Drosophila melanog
gb.in:AC023682	+	378.50	742.78	5.2e-33	179234	AC023682 Drosophila melanog
gb.in:AE003436	+	378.50	737.99	9.6e-33	306135	AE003436 Drosophila melanog
gb.in:AE044601	+	374.50	785.73	2.1e-35	570	AE044601 Sequence 3 from Paten

1720 CTGGCAGCAGACAGACAGTGAACGATTCAACACCTTGGGGAACAT 1769
67 ealaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
1770 TGTCTGACACAGGAGGGTGGCGCAGCCTATAGGCTACCTCAAGTGA 1819
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
1820 TGGCAGAGGTGGCAGGACAGCAGCTGCCGCGCTGGATCTCACCAT 1869
101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114
1870 GAGCAGCTCTTCTCATCACTACGCCAGGTGTGTGGGG 1911

seq_name: gb_pat:AX146980

seq_documentation_block:
LOCUS AX146980 2262 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 5 from Patent WO0136610.
ACCESSION AX146980
VERSION AX146980.1 GI:14346251
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2262)
AUTHORS Deleersnijder, W., Wiegers, R. and Weske, M.
TITLE Human enzymes of the metalloprotease family
JOURNAL Patent: WO 0136610-A 5 25-MAY-2001;
Solvay Pharmaceuticals B.V. (NL)

FEATURES

source
1..2262
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS
1..2262
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC41161.1"
/db_xref="GI:14346251"
/translation="MGKSEGPVGMVESAGRAGKRPGLFEGLLLLLLVLAALVALG
VLVADRRIPEAQEVSEVCTPGCVIAAARTLQNMPTTPCDDDFYQACGWLRRHV
IPETNSRYSIQFVLRDEVLTKAVLENSAKDRPAVEKARTLYRSCMNQSVIEKRG
QPLDILEVGGWVDRNETVGLWELERLALMNSQNRRLVLDLFVNDQNS
SRHIIYIDQPTLGMPSREYFNGSNRKRREAYLQFMVSATLLREDANLPDSCLVQ
EDMMOVLLETQAKATVPOEERHDVIALYHRMGLLELOSOQGLKGFNTLFIQTVLS
SVKILLPDEEVVYGPYQLNLENIIDTYSARTIQNYLVWRLVDRIGLSIQREKDT
RVNYRKALFGTMVEVRWRECVCVYVNSMNAVGLYVREAFPGDSKSMVRELIDKVR
TVFVETLDELGMDEESKKAEKAMSIREQIGHPDYILEEMNRRLDEEYSNLFSD
LYFENSQNLKVGAGRLKRLKREKVDPLNLTIGAAVYNAFSPNRNLQIVFAGILQPP
FFSKEQPOALNFGGIGVIGHEITHFGDNGRDNKNGMMDWNSFNSTOHRFROSEC
MIYQYGNISWDLADEQNVNFTNGENIADNGVRQAYKAYLKWMAEGGKDOPLGLD
LTHEQLPFIYNAQVCGSYRPEFAIQSIKTDVHSPKLYRVLGSLQNLAAAFADTFHAR
GTPMHPKRCRW"

BASE COUNT 520 a 628 c 716 g 398 t
ORIGIN

alignment_scores:
Quality: 651.00 Length: 114
Ratio: 5.711 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-647-780A-4 x AX146980

Align seg 1/1 to: AX146980 from: 1 to: 2262

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17
1756 GGCAGCAGATCAGCAGCGCTTGACGACAATGGCGGAACCTCGACAA 1805
17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34

1806 GAATGGCAACATGATGGATTGGTGGAGTAACCTCTCCACCCAGCACTTCC 1855
34 rgeGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
1856 GGGAGCAGTCAGAGTGCATGATCTACAGTAGCGCACTACTCTCTGGGAC 1905
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsn 67
1906 CTGGCAGACGACAGACAGCTGAACGGATTCAACACCTTGGGGAACAT 1955
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
1956 TGTCTGACAAACGAGGGGTGGCGCAGCCTATAGGCTACCTCAAGTGA 2005
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
2006 TGGCAGAGGTGGCAGGACAGCAGCTGCCGCGCTGGATCTCACCAT 2055
101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114
2056 GAGCAGCTCTTCTCATCACTACGCCAGGTGTGTGGGG 2097

seq_name: gb_pat:AX146978

seq_documentation_block:
LOCUS AX146978 2340 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 3 from Patent WO0136610.
ACCESSION AX146978
VERSION AX146978.1 GI:14346249
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2340)
AUTHORS Deleersnijder, W., Wiegers, R. and Weske, M.
TITLE Human enzymes of the metalloprotease family
JOURNAL Patent: WO 0136610-A 3 25-MAY-2001;
Solvay Pharmaceuticals B.V. (NL)

FEATURES

source
1..2340
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS
1..2340
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC41160.1"
/db_xref="GI:14346250"
/translation="MGKSEGPVGMVESAGRAGKRPGLFEGLLLLLLVLAALVALG
VLVADRRIPEAQEVSEVCTPGCVIAAARTLQNMPTTPCDDDFYQACGWLRRHV
IPETNSRYSIQFVLRDEVLTKAVLENSAKDRPAVEKARTLYRSCMNQSVIEKRG
QPLDILEVGGWVDRNETVGLWELERLALMNSQNRRLVLDLFVNDQNS
SRHIIYIDQPTLGMPSREYFNGSNRKRREAYLQFMVSATLLREDANLPDSCLVQ
EDMMOVLLETQAKATVPOEERHDVIALYHRMGLLELOSOQGLKGFNTLFIQTVLS
SVKILLPDEEVVYGPYQLNLENIIDTYSARTIQNYLVWRLVDRIGLSIQREKDT
RVNYRKALFGTMVEVRWRECVCVYVNSMNAVGLYVREAFPGDSKSMVRELIDKVR
TVFVETLDELGMDEESKKAEKAMSIREQIGHPDYILEEMNRRLDEEYSNLFSD
LYFENSQNLKVGAGRLKRLKREKVDPLNLTIGAAVYNAFSPNRNLQIVFAGILQPP
FFSKEQPOALNFGGIGVIGHEITHFGDNGRDNKNGMMDWNSFNSTOHRFROSEC
MIYQYGNISWDLADEQNVNFTNGENIADNGVRQAYKAYLKWMAEGGKDOPLGLD
LTHEQLPFIYNAQVCGSYRPEFAIQSIKTDVHSPKLYRVLGSLQNLAAAFADTFHAR
GTPMHPKRCRW"

BASE COUNT 539 a 649 c 739 g 413 t
ORIGIN

alignment_scores:
Quality: 651.00 Length: 114
Ratio: 5.711 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-647-780A-4 x AX146978

Align seg 1/1 to: AX146978 from: 1 to: 2340

```
1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAspLy 17
1834 GGCACGAGATCAGCAGCGCTTTCAGCAACAATGGCGGAACCTTCGACAA 1883
17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
1884 GAATGGCAACATGATGGATTGGTGGAGTAACCTCTCCACCCAGCACTTCC 1933
34 rGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
1934 GGGAGCAGTCAGAGTCATGATCTACCATGACGCAACACTCTCTCTGGGAC 1983
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnI 67
1984 CTGGCAGACGACAGAACGTCGACGATTCACACACCTTGGGGAAACAT 2033
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
2034 TGTGCAACAGGAGGCTGCGGCAAGCCTATAGGCTTACCTCAAGTGA 2083
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
2084 TGGCAGAGGTGCGCAAGGACGACGCTGCCGCGCTGGATCTCACCCAT 2133
101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114
2134 GAGCAGCTCTCTTCATCACTACCTACGCGCAGGTGTGTGCGGG 2175
```

seq_name: gb_pat:AX139743

seq_documentation_block:

```
LOCUS AX139743 2636 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 3 from Patent EP1069188.
ACCESSION AX139743
VERSION AX139743.1 GI:14275325
KEYWORDS
```

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2636)

AUTHORS Jagerschmidt A., Agnel M. and Culouscou J.M.

TITLE Three neprilysin-like membrane metalloproteinases

JOURNAL Patent: EP 1069188-A 3 17-JAN-2001;

SANOFI-SYNTHELABO (FR)

Location/Qualifiers

1. .2636

/organism="Homo sapiens"

/db_xref="taxon:9606"

1. .7

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC39975.1"

/db_xref="GI:14275326"

```
/translation="MGKSEGPVGMVESAGRAGKRPGLFEGGLLLLLLVTAALVALG
VLADRRIPEAQEYSEVCTPGCVIAARILQNDPTTEPCDDFYQFACGWLRRHY
IPETNSRYSIFDLVLEILKAVLENSTAKDRPAVEKARTLYRSCNQSIVERRGS
OPLLDILEVGGWPVAMDNRNFTVGLWELEQLALMNSFNRRVLDLFIWDDONS
SRHIYIDOPTLGMPSREYFNGSGNRKREAYLOFMSVATLLREDANLPRDCLVQ
EDMVQLELETQALATVPEERHVDIALYHRMGLEELQSQGLKGFDTWLTFTVLS
SVKILPDVEYVYVIGIPYLNLENLIDYTSARTIQNYLWRLDLRGLSLSRFKDT
RYNKRKALFTVVEVRRECYGVYNSNMENAVSLYREAFPGDSKSMVRELIDKVR
TVFETLDELGMDEESKKAEKAMSIREQIGHDPYILEETNRLDEESYLNLFSED
LYFENSLONLKVCAORSRLKREKVDPNLWIIIGAAVFNAYFSPNRNQIVFPAGILQPP
FFSKOPQALNFGGIMWICHTGHTGFDNGRNFKNMGMDWNSFNSTQHFREQSEC
MIYQYGNYSWDLADQSNVNGFNTLGENIADNGVROAYKAYLKWMAEGGKDDQLGLD
LTHEQLFTNYAQVWCGSYRPEFAIOSIKTVDHSPLKRYVLGSLQNLAAFAADTFHCAR
GTPMHPKRCRVW"
```

FEATURES

source

5'UTR

CDS

```
3'UTR 2270: .2635
BASE COUNT 601 a 759 c 813 g 462 t 1 others
ORIGIN
```

```
alignment_scores:
  Quality: 651.00 Length: 114
  Ratio: 5.711 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

US-09-647-780A-4 x AX139743

Align seg 1/1 to: AX139743 from: 1 to: 2636

```
1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAspLy 17
1763 GGCACGAGATCAGCAGCGCTTTCAGCAACAATGGCGGAACCTTCGACAA 1812
17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
1813 GAATGGCAACATGATGGATTGGTGGAGTAACCTCTCCACCCAGCACTTCC 1862
34 rGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
1863 GGGAGCAGTCAGAGTCATGATCTACCATGACGCAACACTCTCTCTGGGAC 1912
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnI 67
1913 CTGGCAGACGACAGAACGTCGACGATTCACACACCTTGGGGAAACAT 1962
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
1963 TGTGCAACAGGAGGCTGCGGCAAGCCTATAGGCTTACCTCAAGTGA 2012
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
2013 TGGCAGAGGTGCGCAAGGACGACGCTGCCGCGCTGGATCTCACCCAT 2062
101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114
2063 GAGCAGCTCTCTTCATCACTACCTACGCGCAGGTGTGTGCGGG 2104
```

seq_name: gb_pat:AX139745

seq_documentation_block:

```
LOCUS AX139745 2663 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 5 from Patent EP1069188.
ACCESSION AX139745
VERSION AX139745.1 GI:14275327
KEYWORDS
```

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2663)

AUTHORS Jagerschmidt A., Agnel M. and Culouscou J.M.

TITLE Three neprilysin-like membrane metalloproteinases

JOURNAL Patent: EP 1069188-A 5 17-JAN-2001;

SANOFI-SYNTHELABO (FR)

Location/Qualifiers

1. .2663

/organism="Homo sapiens"

/db_xref="taxon:9606"

1. .7

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC40013.1"

/db_xref="GI:14275328"

```
/translation="MGKSEGPVGMVESAGRAGKRPGLFEGGLLLLLLVTAALVALG
VLADRRIKQLPLASRLFLQERTFVKRPKGPICPAQSEVCTTPGCVIAARIL
QNMDPTTEPCDDFYQFACGWLRRHYIPETNSRYSIFDLVLEILKAVLENSTAK
```

DRPVEKARTLYRSCNMOSVIEKRSOPLLDILEVGGWVPVMDRWNETVVQARPHR
PLHREPELOPAHLLHPTLGMPSREYFNGSGNRKVRAYLOFMVSATLLREDAN
LPDSCLYQEDMVOVLEETQLAKATVPOEERHDVIALYHRMGLEELQSGFGLKGFDM
TLFTQTVLSSVKIKLPDEEVVVYIGYILQNLNENIDITYSARTTIONLYVRLVLDRI
SLSQKDFRTVNTKALFGTMVEVRRECQVYVNSNNAAGSLYVREAFPGDKSM
VRELIDKRVTVFVETLDELGMDESKKKAQERAMSREOITGHDPDYILEETNRRLDEE
YSNLISSEDLFFENSLONLKVGARSURKLRKVDPLNWIIGAAVNAFYSPNRNOITV
FPAGILQPPFSKQOPALNFGGIMVIGHEITHGFDGNGRNFDKNGMMDWSNFT
OHPREOSECMIYQYGNYSWDLAEOVNGENTLGENIADNGGVROAKAYLKNWAEGL
KQOOLPGDLTHTEOLFINTAQVWCYSIRPEFAIOSIKTDVHSPKRYVLGSLQNLAA
FADTFHCARTPMHPKRCRW"
2294 .2663
3' UTR 609 a 769 c 816 g 468 t 1 others
BASE COUNT 608 a 771 c 823 g 474 t
ORIGIN

alignment_scores:
Quality: 651.00 Length: 114
Ratio: 5.711 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-647-780A-4 x AX139745

Align seg 1/1 to: AX139745 from: 1 to: 2663

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17
|||||
1790 GGGCAGCAGATCAGCAGCGCTTTGACGCAATGGCCGGAACCTTCGACAA 1839

17 sAsnGlyAsnMetMetAspTrpSerAsnPheSerThrGlnHisPheA 34
|||||
1840 GAATGGCACATGATGGATTGGTGAGTAGTACTTCTCCACCCAGCAGCTTC 1889

34 rGluGlnSerGluCysMetIleTyrglnTyrglyAsnTyrSerTrpAsp 50
|||||
1890 GGGAGCAGTCAGAGTCATGATCTACCACTAGCGCAACTACTCTCTGGGAC 1939

51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsn 67
|||||
1940 CTGGCAGACGACAGACAGTGAACGGATTCAACACCCCTTGGGGAAACAT 1989

67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrp 84
|||||
1990 TGCTACACAGGAGGGGTCGCGCAAGCCTTAAGGCTTACCTCAAGTGA 2039

84 eTalaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
|||||
2040 TGGCAGAGGGTGGCAAGGACAGCAGCTGCCCGCTGGATCTCACCCAT 2089

101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114
|||||
2090 GAGCAGCTCTTCTCATCACTATGCCAGGTCGCGGTGGTGGCGG 2131

seq_name: gb_pat:AX033274

seq_documentation_block:
LOCUS AX033274 2676 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 14 from Patent W00047750.
ACCESSION AX033274
VERSION AX033274.1 GI:10280089.

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2676)

AUTHORS Boileau, G. and Desgrosseillers, L.

TITLE New metalloproteases of the neprilysin family

JOURNAL Patent: WO 0047750-A 14 17-AUG-2000;

BOILEAU GUY (CA) ; DESGROSSEILLERS LUC (CA) ; UNIVERSITE DE MONTREAL

(CA)

FEATURES Location/Qualifiers

source 1. 2676
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 7. 2319
/note="unnamed protein product".
/codon_start=1
/protein_id="CAC09978.1"
/db_xref="GI:10280090"

/translation="MVESAGRAGOKRPGFLEGGLLLLLLVTAALVALGVLYADRRGK
OLPLRLAKRLQBERTFVRKRKPGIPPAQSEVCTTPGCVIAAARILQNMDDPTPEP
CDDPYACGGWLRHRVITPNSRYSIDFVLRDELEVLKAVLENSTAKDRPAVBKAR
TLYRSCNMOSVIEKRSOPLLDILEVGGWVPVMDRWNETVGLMELEBROLALMNSQF
NRRVILDLFIWDDONSRRHIIYIDOPTLGMPSREYFNGSGNRKVRAYLOFMVSVA
TLREDANLPDSCLYQEDMVOVLEETQLAKATVPOEERHDVIALYHRMGLEELQSG
FGLKGFNTLFIQTVLSSVKIKLPDEEVVVYIGYILQNLNENIDITYSARTTIONLYV
RLVLDRIQSGREFKTRVNTKALFGTMVEVRRECQVYVNSNNAAGSLYVREAFPGDKSM
VRELIDKRVTVFVETLDELGMDESKKKAQERAMSREOITGHDPDYILEETNRRLDEE
YSNLISSEDLFFENSLONLKVGARSURKLRKVDPLNWIIGAAVNAFYSPNRNOITV
FPAGILQPPFSKQOPALNFGGIMVIGHEITHGFDGNGRNFDKNGMMDWSNFT
OHPREOSECMIYQYGNYSWDLAEOVNGENTLGENIADNGGVROAKAYLKNWAEGL
KQOOLPGDLTHTEOLFINTAQVWCYSIRPEFAIOSIKTDVHSPKRYVLGSLQNLAA
FADTFHCARTPMHPKRCRW"
BASE COUNT 608 a 771 c 823 g 474 t
ORIGIN

alignment_scores:
Quality: 651.00 Length: 114
Ratio: 5.711 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-647-780A-4 x AX033274

Align seg 1/1 to: AX033274 from: 1 to: 2676

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17
|||||
1813 GGGCAGCAGATCAGCAGCGCTTTGACGCAATGGCCGGAACCTTCGACAA 1862

17 sAsnGlyAsnMetMetAspTrpSerAsnPheSerThrGlnHisPheA 34
|||||
1863 GAATGGCACATGATGGATTGGTGAGTAGTACTTCTCCACCCAGCAGCTTC 1912

34 rGluGlnSerGluCysMetIleTyrglnTyrglyAsnTyrSerTrpAsp 50
|||||
1913 GGGAGCAGTCAGAGTCATGATCTACCACTAGCGCAACTACTCTCTGGGAC 1962

51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsn 67
|||||
1963 CTGGCAGACGACAGACAGTGAACGGATTCAACACCCCTTGGGGAAACAT 2012

67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrp 84
|||||
2013 TGCTGACACAGGAGGGTGGCGCAAGCCTTAAGGCTTACCTCAAGTGA 2062

84 eTalaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
|||||
2063 TGGCAGAGGGTGGCAAGGACAGCAGCTGCCCGCTGGATCTCACCCAT 2112

101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114
|||||
2113 GAGCAGCTCTTCTCATCACTATGCCAGGTCGCGGTGGTGGCGG 2154

seq_name: gb_pat:AX139741

seq_documentation_block:

LOCUS AX139741 2714 bp DNA linear PAT 30-MAY-2001

DEFINITION Sequence 1 from Patent EP1069188.

ACCESSION AX139741

VERSION AX139741.1 GI:14275323

KEYWORDS

SOURCE human.

```
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2714)
AUTHORS Jagerschmidt,A., Agnel,M. and Culoussou,J.M.
TITLE Three neprilysin-like membrane metalloproteinases
JOURNAL Patent: Ep 1069188-A 1 17-JAN-2001;
SANOFI-SYNTHELABO (FR)
FEATURES
    source      Location/Qualifiers
                1. 2714
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                5'UTR
                CDS
                    1. 2347
                    /note="unnamed protein product"
                    /protein_id="CAC39974.1"
                    /db_xref="GI:14275324"
                    /translation="MGKSEGPVGMVESAGRAGKRPGLGGLLLLLLVTAALVALG
                    VLYADRGKQLPLASFLCLPLOBERTFVKPKPGIPEAQEVSEVCTTPGCVIAAARIL
                    DRPAVEKARTLYRSCNNQSVIEKRGSQPLDILEVGGVPMVDRWNETVGLWELEL
                    QLALMNSQNRRLVLDIFWINDQNSRHIIYIDQTLGMPSEYFYFGSGNRKVR
                    YLQPMVSATLLKGFDMFTLFTQTLSSVKIKLLPDEEVVYGIPLYQNLNIDITYA
                    RTQNYLVRLVLDRLGSLRSQRFKTRVNYRKALFGTMVEVRKRECVGVNSNWE
                    VGLSYVREAPFGDSKSMRELIDKRVTVFETLDELGMDEESKKKAEKAMSTREQI
                    GHPTDILEETNRRLDEYSNLFNSEDLYFNSLQNLKVGARSUKRLKRVDPNLWII
                    GAAVNAFYSNPNRQIVFPAGIQLPPFSEKQPOALNFGGIGMWIGHEITHGFDDNG
                    NFDKNGNMWDSNFSFTHREQSECMYQYGNYSWDLADQNQVNGFNTLGENADNG
                    GVPOAYKYLKWAEGKDOOLPGLDLTHEQLFFINVAQVWCGSYRPEFAIOSIKTDV
                    HSPKLYVLSGLNLAFAFTHCARGTPMHPKRCRW"
                3'UTR
                2348..2713
                BASE COUNT      621 a 780 c 836 g 476 t      1 others
                ORIGIN

alignment_scores:
    Quality: 651.00      Length: 114
    Ratio: 5.711        Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-647-780A-4 x AX139741
Align seg 1/1 to: AX139741 from: 1 to: 2714
1 GlyHisGluIleThrHisGlyPheAspAspAsnGlyArgAsnPheAsp 17
1841 GGGCAGAGATCAGCAGCGCTTTGACGAAATGGCGGGAACATTCGACAA 1890
17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
1891 GAATGCAACATGATGATGGTGGGAGTAACTTCTCCACCACGACACTTC 1940
34 rgGluGlnSerGluCysMetIleTyrglnTyrglyAsnTyrsTrpAsp 50
1941 GGGAGCAGTCAGAGTGCATGATCATTACAGTACGCAACTACTCTCTGGGAC 1990
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnII 67
1991 CTGGCAGACGACAGACAGCGTGAACGGATTCAACACCCCTTGGGAAACAT 2040
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrlLysAlaTyrlLeuLysTrpM 84
2041 TGCTGCAACGAGGAGGGTGGCGAAGCCTTAAAGCCTACTCTCAAGTGA 2090
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
2091 TGGCAGAGGTGGCAAGACGACGAGCTGCCCGGCTGGATCTCACCCAT 2140
101 GluGlnLeuPhePheIleAsnTyrlAlaGlnValTrpCysGly 114
|||||
```

```
2141 GAGCAGCTCTTCTTCATCAACTATGCCCGAGGTGTGTGTCGGG 2182
seq_name: gb_pr:AF336981
seq_documentation_block:
LOCUS AF336981 2784 bp mRNA linear PRI 01-OCT-2001
DEFINITION Homo sapiens neprilysin-like metalloproteinase 2 mRNA, complete cds.
ACCESSION AF336981
VERSION AF336981.1 GI:15811370
KEYWORDS
SOURCE
    human.
    ORGANISM Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    REFERENCE 1 (bases 1 to 2784)
    AUTHORS Bonvouloir,N., Lemieux,N., Crine,P., Boileau,G. and
        DesGroselliers,L.
    TITLE Molecular cloning, tissue distribution, and chromosomal
        localization of MMEL2, a gene coding for a novel human member of
        the neutral endopeptidase-24.11 family
    JOURNAL DNA Cell Biol. 20 (8), 493-498 (2001)
    MEDLINE 21444797
    PUBMED 11560781
    REFERENCE 2 (bases 1 to 2784)
    AUTHORS Bonvouloir,N., Lemieux,N., Crine,P., Boileau,G. and
        DesGroselliers,L.
    TITLE Direct Submission
    JOURNAL Submitted (15-JAN-2001) Biochemistry, University of Montreal, P. O.
        Box 6128, Station Centre Ville, Montreal, Quebec H3C 3J7, Canada
    FEATURES
        source
            1. 2784
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /map="lp36"
                /chromosome="1"
                /note="NEP-like 2; zinc metalloproteinase"
                /codon_start=1
                /product="neprilysin-like metalloproteinase 2"
                /protein_id="AAI08942.1"
                /db_xref="GI:15811371"
                /translation="MVESAGRAGKRPGLGGLLLLLLVTAALVALGVLVADRRGK
                OLPLASRLCFLOBERTFVKPKPGIPEAQEVSEVCTTPGCVIAAARILQNDMDPTPEP
                CDDFYQACGGWLRHRVLPETNSRYSIFDLRDELEVILKVALENSTAKDRPAVEKAR
                TLYRSCNNQSVIEKGSQPLDILEVGGVPMVDRWNETVGLWELEQLALMNSQV
                NRVLDLFINDDQNSRHIIYIDQTLGMPSEYFYFGSGNRKVRKREAYLQPMVSA
                TLLREDANLPDRSCLVQDMVQVLETLQALATVPQERHDVIALYHRMGLEELQSQ
                FGLKGNMTLFTQTLSSVKIKLLPDEEVVYGIPLYQNLNIDITYSARTQNYLVW
                RLVLDRIGLSQRFKTRVNYRKALFGTMVEVRKRECVGVNSNMENAVGSLYVREA
                PPGDSKSMRELIDKRVTVFETLDELGMDEESKKKAEKAMSTREQIIGHPDYILEE
                MNRLDDEYSNLFNSEDLYFNSLQNLKVGARSUKRLKREKVDPNLWIIIGAAVVAFY
                SPNRNQIVFPAGIQLPPFSEKQPOALNFGGIGMWIGHEITHGFDDNGRNFKNMNM
                DWMSFSTQHFREQSECMYQYGNYSWDLADQNQVNGFNTLGENIADNGSGVRQAYKAY
                LKWAEGKDOOLPGLDLTHEQLFFINVAQVWCGSYRPEFAIOSIKTDVHSPKLYRVL
                GSLNLAFAFTHCARGTPMHPKRCRW"
        BASE COUNT      669 a 789 c 841 g 485 t
        ORIGIN

alignment_scores:
    Quality: 651.00      Length: 114
    Ratio: 5.711        Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
US-09-647-780A-4 x AF336981
Align seg 1/1 to: AF336981 from: 1 to: 2784
1 GlyHisGluIleThrHisGlyPheAspAspAsnGlyArgAsnPheAsp 17
|||||
1871 GGGCAGAGATCAGCAGCGCTTTGACGACAAATGCCCGGAACATTCGACAA 1920
```



```
17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
|||||
1921 GAATGGCAACATGATGATTGGTGGAGTAACCTTCCACCCAGCACTTCC 1970

34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpasp 50
|||||
1971 GGGAGCAGTTCAGAGTGCATGATCTACCAAGTACGGCAACTACTCTCTGGGAC 2020

51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnI1 67
|||||
2021 CTGGCAGCAGGAGACAGACGAGTGCAGGATTCACACCTTGGGGAAACAT 2070

67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
|||||
2071 TGCTGACAACGGAGGGGTGCGCAAGCTATAAGGCTTACCTCAAGTGA 2120

84 eTalaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
|||||
2121 TGGCAGAGGGTGGCAAGACAGCAGCTGCCCGCTGGATCTCACCCAT 2170

101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTyrCysGly 114
|||||
2171 GAGCAGCTCTCTTCATCACTATGCCAGGTGTGTGGCGG 2212

seq_name: gb_pat:AX319864
seq_documentation_block:
LOCUS AX319864 2232 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 28 from Patent WO0183782.
ACCESSION AX319864
VERSION AX319864.1 GI:17901454
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Plowman,G.D., Whyte,D., Sudarsanam,S., Manning,G., Caenepeel,S. and
Payne,V.
TITLE Novel proteases
JOURNAL Patent: WO 0183782-A 28 08-NOV-2001;
Sugen, Inc. (US).
FEATURES
source
1..2232
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 512 a 620 c 705 g 395 t
ORIGIN
alignment_scores:
Quality: 640.50 Length: 115
Ratio: 5.618 Gaps: 1
Percent Similarity: 99.130 Percent Identity: 99.130
alignment_block:
US-09-647-780A-4 x AX319864
Align seg 1/1 to: AX319864 from: 1 to: 2232
.1 GlyHisGluIleThrHisGlyPheAspAsn...GlyArgAsnPheAs 16
|||||
1723 GGCACAGATCAGCAGCGCTTTCAGCACATGGTGGCGGAACTTCGA 1772

16 pLysAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisP 33
|||||
1773 CAAGAATGGCAACATGATGGTGGTGGAGTAACCTTCCACCCAGCACT 1822

33 heargGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrp 49
|||||
1823 TCCGGGAGCAGTTCAGAGTGCATGATCTACCAAGTACGGCAACTACTCTCG 1872

50 AspLeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAs 66
|||||
```

```
1873 GACCTGGCAGACGACAGACGATGAACGATTCACACCCTTGGGGAAA 1922

66 nIleAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysT 83
|||||
1923 CATTCGTGACAACGGAGGGGTGCGCAAGCTATAAGGCTTACCTCAAGT 1972

83 rPmetAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThr 99
|||||
1973 GGATGGCAGAGGGTGGCAAGACAGCAGCTGCCCGCTGGATCTCACC 2022

100 HisGluGlnLeuPhePheIleAsnTyrAlaGlnValTyrCysGly 114
|||||
2023 CATGACGAGCTCTCTTCATCACTATGCCAGGTGTGTGGCGG 2067

seq_name: gb_pat:AX014703
seq_documentation_block:
LOCUS AX014703 327 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 3 from Patent WO9953077.
ACCESSION AX014703
VERSION AX014703.1 GI:10040977
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 327)
AUTHORS Schwartz,J.C., Gros,C., Ouimet,T., Rose,C., Bonhomme,M.C. and
Facchinetti,P.
TITLE Novel nep ii membrane metalloprotease and its use for screening
inhibitors useful in therapy
JOURNAL Patent: WO 9953077-A 3 21-OCT-1999;
INST NAT SANTE RECH MED (FR); SCHWARTZ JEAN CHARLES (FR); GROS
CLAUDE (FR); OUIMET TANJA (FR); ROSE CHRISTIANE (FR); BONHOMME
MARIE CHANTAL (FR); FACCHINETTI PATRICIA (FR)
FEATURES
source
1..327
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 87 a 88 c 93 g 59 t
ORIGIN
alignment_scores:
Quality: 616.00 Length: 109
Ratio: 5.651 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-647-780A-4 x AX014703
Align seg 1/1 to: AX014703 from: 1 to: 327
1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17
|||||
1 GGCACAGATCAGCAGCGCTTTCAGCACATGGTGGCGGAACTTCGACAA 50

17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
|||||
51 GAATGGCAACATGATGATGGTGGTGGAGTAACCTTCCACCCAGCACTTCC 100

34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpasp 50
|||||
101 GGGAGCAGTTCAGAGTGCATGATCTACCAAGTACGGCAACTACTCTCTGGGAC 150

51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnI1 67
|||||
151 CTGGCAGCAGGAGACAGACGATTCACACCTTGGGGAAACAT 200

67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
|||||
201 TGCTGACAACGGAGGGGTGCGCAAGCTATAAGGCTTACCTCAAGTGA 250
```

84 etAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
 |||||
 251 TGGCAGAGGGTGGCAAGCAGCAGCAGTGGCCGGCTGGATCTACCCAT 300
 |||||

101 GluGlnLeuPhePheIleAsnTyrAla 109
 |||||
 301 GAGCAGCTCTTCTTCATCACTATGCC 327
 |||||

seq_name: gb_ro:AF302075

seq_documentation_block:

LOCUS AF302075 2583 bp mRNA linear ROD 11-JUN-2001

DEFINITION Mus musculus neprilysin-like peptidase alpha mRNA, complete cds.

ACCESSION AF302075

VERSION AF302075.1 GI:10505359

SOURCE

house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Shirotani, K., Tsubuki, S., Iwata, N., Takaki, Y., Harigaya, W.,

Maruyama, K., Kiryu-Seo, S., Kiyama, H., Iwata, H., Tomita, T.,

Iwatsubo, T. and Saido, T.C.

Neprilysin Degrades Both Amyloid beta Peptides 1-40 and 1-42. Most

Rapidly and Efficiently among Thiorphan- and

Phosphoramidon-sensitive Endopeptidases

J. Biol. Chem. 276 (24), 21895-21901 (2001).

JOURNAL

PUBLISHED

REFERENCE

AUTHORS

Shirotani, K. and Saido, T.C.

Direct Submission

TITLE

Submitted (02-SEP-2000) Proteolytic Neuroscience Laboratory, Brain

Science Institute, RIKEN, 2-1 Hirotsawa, Wako-shi, Saitama

351-0198, Japan

FEATURES

source

1. 2583

/organism="Mus musculus"

/db_xref="taxon:10090"

25. 2253

/note="endopeptidase"

/codon_start=1

/product="neprilysin-like peptidase alpha"

/protein_id="AAG18446.1"

/db_xref="GI:10505360"

/translation="MVERAGWCRKSPGFVEYGLMVLALLLGAIVTLGVFYSIGKQL

SSLSKSDICTTPSCVIAAARTLENDDSPNCFYQVACGWLRRHHVPTNRSYVF

DILRDEVLTKGLVEDSTSOHRPAVEKATLYRSCMNQSVIEKRDSEPLLSVLMVG

GHPVAMDKNETHGLKHELRLQAVLNSQNRRLVLDLFTWDDQNSRRHVYIDQPT

LGMSREYFQEDNNHKKRKYLEFMTSVATMLRKDQNLKSAMVREMAEVLLEET

HLANATVPOEKRDHVTALYHRMDLMEQLQERGLKFNWTLFIQNLVSVEVLPDDE

VVYGIPILENLEDDIIDSARTMQLYVWRLVLDRLIGLSQRFKEARVDYRKALYT

TVEVRWRECVSVNMSAVGLYIKRAFSDKSTVRELIEKRSVFVDNLDELN

WDEESKKAKQAKMIREIGYDPVILEDNKKHLDDEYSLAFYEDLYFENGIONLK

NAGRSRLKDKREYDQNLWIIIGAAYVNAFVSPNRNQLVFPAGILOPFFSKDQPSLN

FGGIVGVIHREYTHGDDNDFDKNGMLDWSNFARFQOQSCMCIYOYGNFWE

LADQNVNGFSTLGENIADNGVRQAYKAYLRWLADGDKDQRLPGLNLTYAQLFFNY

AQVWCYSRFEFAVQSIKTDVHSPKLYRVLGSLQNLPCFSEAFHCPRGSPMHPMKRCR

IW"

BASE COUNT 665 a 667 c 736 g 515 t

ORIGIN

alignment_scores:

Quality: 581.00 Length: 114

Ratio: 5.142 Caps: 0

Percent Similarity: 99.123 Percent Identity: 85.088

alignment_block:

US-09-647-780a-4 x AF302075

Align seg 1/1 to: AF302075 from: 1 to: 2583

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPhesply 17
 |||||
 1747 GGGCAGAGATCACACACGGCTTTGATGATATGTCGTAACCTTTGACAA 1796
 |||||

17 sasnglyasnmetmetaspttrpserasnpheserthrGlnHisPhea 34
 |||||
 1797 GAAACGGCAACATGCTGGACTGGTGGAGTAACCTCTCGGCCCGCACTCC 1846
 |||||

34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
 |||||
 1847 AACAGCAGTCGCAATGCAATGATCTATCAGTACGGCAACTTCTCTTGGAA 1896
 |||||

51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsn11 67
 |||||
 1897 CTAGCAGACACCAACAGATGTAACGGATTCAGTACCTCGGGGAGACAT 1946
 |||||

67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
 |||||
 1947 TGCCGACACAGCGAGGTGTCGACAGGCAATACAAAGGCTTACCTACGGTGGC 1996
 |||||

84 etAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
 |||||
 1997 TGGCTGATGGCGGCAAGATCAGCGACTCCCGGAGCTGAACCTGACCTAT 2046
 |||||

101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114
 |||||
 2047 GCCCAGCTTTTCTTATCAATATGCCAGAGGTGGTGTGGG 2088
 |||||

seq_name: gb_ro:AF302076

seq_documentation_block:

LOCUS AF302076 2652 bp mRNA linear ROD 11-JUN-2001

DEFINITION Mus musculus neprilysin-like peptidase beta mRNA, complete cds.

ACCESSION AF302076

VERSION AF302076.1 GI:10505361

KEYWORDS

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Shirotani, K., Tsubuki, S., Iwata, N., Takaki, Y., Harigaya, W.,

Maruyama, K., Kiryu-Seo, S., Kiyama, H., Iwata, H., Tomita, T.,

Iwatsubo, T. and Saido, T.C.

Neprilysin Degrades Both Amyloid beta Peptides 1-40 and 1-42. Most

Rapidly and Efficiently among Thiorphan- and

Phosphoramidon-sensitive Endopeptidases

J. Biol. Chem. 276 (24), 21895-21901 (2001)

JOURNAL

PUBLISHED

REFERENCE

AUTHORS

Shirotani, K. and Saido, T.C.

Direct Submission

TITLE

Submitted (02-SEP-2000) Proteolytic Neuroscience Laboratory, Brain

Science Institute, RIKEN, 2-1 Hirotsawa, Wako-shi, Saitama

351-0198, Japan

FEATURES

source

1. 2652

/organism="Mus musculus"

/db_xref="taxon:10090"

25. 2322

/note="endopeptidase"

/codon_start=1

/product="neprilysin-like peptidase beta"

/protein_id="AAG18447.1"

/db_xref="GI:10505362"

/translation="MVERAGWCRKSPGFVEYGLMVLALLLGAIVTLGVFYSIGKQL

PLLSLHLFSDERTVVRALRDLSSLSKSDICTTPSCVIAAARTLENDDSPNCFY

QYACGGWLRRHHVPTNRSYVFELRDEVLTKGLVEDSTSOHRPAVEKATLYRSCMNQSVIEKRDSEPLLSVLMVG

GHPVAMDKNETHGLKHELRLQAVLNSQNRRLVLDLFTWDDQNSRRHVYIDQPT

LGMSREYFQEDNNHKKRKYLEFMTSVATMLRKDQNLKSAMVREMAEVLLEET

HLANATVPOEKRDHVTALYHRMDLMEQLQERGLKFNWTLFIQNLVSVEVLPDDE

VVYGIPILENLEDDIIDSARTMQLYVWRLVLDRLIGLSQRFKEARVDYRKALYT

TVEVRWRECVSVNMSAVGLYIKRAFSDKSTVRELIEKRSVFVDNLDELN

WDEESKKAKQAKMIREIGYDPVILEDNKKHLDDEYSLAFYEDLYFENGIONLK

NAGRSRLKDKREYDQNLWIIIGAAYVNAFVSPNRNQLVFPAGILOPFFSKDQPSLN

FGGIVGVIHREYTHGDDNDFDKNGMLDWSNFARFQOQSCMCIYOYGNFWE

LADQNVNGFSTLGENIADNGVRQAYKAYLRWLADGDKDQRLPGLNLTYAQLFFNY

AQVWCYSRFEFAVQSIKTDVHSPKLYRVLGSLQNLPCFSEAFHCPRGSPMHPMKRCR

IW"

BASE COUNT 665 a 667 c 736 g 515 t

ORIGIN

alignment_scores:

Quality: 581.00 Length: 114

Ratio: 5.142 Caps: 0

Percent Similarity: 99.123 Percent Identity: 85.088

alignment_block:

US-09-647-780a-4 x AF302075

Align seg 1/1 to: AF302075 from: 1 to: 2583

RIGLSQREKARVDYKALYGTVEEVRWRCVSVVNSNMESAVGLYIKRAFSDS
KTVRELIEKIRSVFVNDLDELNMDEESKKAEKAMNIREQIGYDPDYLIDNNKHL
DEYSLEYEDLYFVNDLQNNKNAOSIKLREKVDQNLHIGAAVNVNAYSPNRN
QIVFAGILQPPFFSKDQPSNFGIGWIGHEITHGDDNGRDNFKNGMLDWSN
FSARHFQOQSCMIYOYGNFSEWELADQNVNFGFSTIGENIADNGGVQAYKAYLRWLA
DGKQORLGLNLTAYALFFINAYAOVWCGSYRPEFAVOSIKTDVHSPLKYRVLGSLQ
LPGFSEAFHCPRGSPHMKRCRIW"
BASE COUNT 682 a 685 c 755 g 530 t
ORIGIN

alignment_scores:
Quality: 581.00 Length: 114
Ratio: 5.142 Gaps: 0
Percent Similarity: 99.123 Percent Identity: 85.088

alignment_block:

US-09-647-780A-4 x AF302076

Align seg 1/1 to: AF302076 from: 1 to: 2652

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17

1816 GGCACGAGATCACACACGGCTTTCATGATATGCTCGTAACCTTGACAA 1865

17 sAsnGlyAsnMetMetAspTtTpSerAsnPheSerThrGlnHisPheA 34

1866 GAACGGCAACATGCTGGACTGGTGAGTAACCTTCGCGCCCGGCACCTTCC 1915

34 rGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50

1916 AACAGCAGTCGCAATGATGATCATCATGACGACGCACTTCTCTGGGAA 1965

51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsn 67

1966 CTAGCAGACACACAGAAATGTGAACGGATTACCTACCTCGCGGGAGAACAT 2015

67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84

2016 TGCCGACACGGAGGTGTGCGACAGGCATACAAAGCTTACCTACGCTGGC 2065

84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100

2066 TGGCTGATGGCGGCAAGATCAGCGACTGCGCGGACTGAACCTGACCTAT 2115

101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114

2116 GCCCAGCTTTTCTTCATCACTATGCCAGGTGTGGTGGG 2157

seq_name: gb_ro:AF302077

seq_documentation_block:

LOCUS AF302077 2694 bp mRNA linear ROD 11-JUN-2001
DEFINITION Mus musculus neprilysin-like peptidase gamma mRNA, complete cds.

ACCESSION AF302077

VERSION AF302077.1 GI:10505363

KEYWORDS

SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 2694)

Shirohani, K., Tsubuki, S., Iwata, N., Takaki, Y., Harigaya, W.,

Mariyama, K., Kiryu-Seo, S., Kiyama, H., Iwata, H., Tomita, T.,

Iwatsubo, T. and Saido, T.C.

Neprilysin Degrades Both Amyloid beta Peptides 1-40 and 1-42 Most

Rapidly and Efficiently among Thiorphan- and

Phosphoramidon-sensitive Endopeptidases

J. Biol. Chem. 276 (24), 21895-21901 (2001)

11278416

2 (bases 1 to 2694)

Shirohani, K. and Saido, T.C.

Direct Submission

JOURNAL

Submitted (02-SEP-2000) Proteolytic Neuroscience Laboratory, Brain
Science Institute, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama
351-0198, Japan

FEATURES

source

1. 2694

/organism="Mus musculus"

/db_xref="taxon:10090"

25..2364

/note="endopeptidase"

/codon_start=1

/product="neprilysin-like peptidase gamma"

/protein_id="AAG18448.1"

/db_xref="GI:10505364"

/translation="MVERAGCRKKSPFVEYGLMVLALLLLGAVTLGVFYSIALRD
SSLKSDICTPSCVIAAARILLENDOSRPNCFYOVACGMLRHHVLPETNSRYSVF
DILRDELEVLKGVLEDTSQHRPAVEKAKTLYRSCNNQSVIEKRDSPELLSVLKMVG
GWPVAMKWNETMGLWELEKQALVNSQNRRLRIDLFINWDDQNSRHHVYIDQPT
HLANATYPOEKRHDTALYHRMDLMELQERGLKDRVSLCSPGCGTHSDVQAGLELG
NPASDSRVLGKGFNNTLFIQNVLSSEVELEFDEEVVYGYPILENLEIDISYA
RTMONTLVRLVLDRLGSLRSQREKARVDYKALYGTVEEVRWRCVSVVNSNMESA
VGSLYIKRAFSDKSDSTVRELIEKIRSVFVNDLDELNMDEESKKAEKAMNIREQI
GYDYLIDNNKHLDEYSLEYEDLYFVNDLQNNKNAOSIKLREKVDQNLHIGAAVNVN
AAVNAFYSPNRNIVFPAGILOPPFFSKDQPSNFGIGWIGHEITHGDDNGRDNFKNG
NFDKNGMLDWSNFSARHFQOQSCMIYOYGNFSEWELADQNVNFGFSTIGENIADNG
GVROAYKAYLRWLADGKQORLGLNLTAYALFFINAYAOVWCGSYRPEFAVOSIKTDV
HSPFKYRVLGSLQNLPGFSEAFHCPRGSPHMKRCRIW"

BASE COUNT 686 a 700 c 766 g 542 t

ORIGIN

alignment_scores:

Quality: 581.00 Length: 114

Ratio: 5.142 Gaps: 0

Percent Similarity: 99.123 Percent Identity: 85.088

alignment_block:

US-09-647-780A-4 x AF302077

Align seg 1/1 to: AF302077 from: 1 to: 2694

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17

1858 GGCACGAGATCACACACGGCTTTCATGATATGCTCGTAACCTTGACAA 1907

17 sAsnGlyAsnMetMetAspTtTpSerAsnPheSerThrGlnHisPheA 34

1908 GAACGGCAACATGCTGGACTGGTGAGTAACCTTCGCGCCCGGCACCTCC 1957

34 rGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50

1958 AACAGCAGTCGCAATGATGATCATCATGACGCACTTCTCTGGGAA 2007

51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsn 67

2008 CTAGCAGACACACAGAAATGTGAACGGATTACCTACCTCGCGGGAGAACAT 2057

67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84

2058 TGCCGACACGGAGGTGTGCGACAGGCATACAAAGCTTACCTACGCTGGC 2107

84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100

2108 TGGCTGATGGCGGCAAGATCAGCGACTGCGCGGACTGAACCTGACCTAT 2157

101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114

2158 GCCCAGCTTTTCTTCATCACTATGCCAGGTGTGGTGGG 2199

seq_name: gb_pat:AX033272

seq_documentation_block:

LOCUS AX033272

2925 bp

DNA

linear

PAT 21-SEP-2000

```

DEFINITION Sequence 12 from patent WO0047750.
ACCESSION AX033272
VERSION AX033272.1 GI:10280087
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2925)
AUTHORS Boileau,G. and Desgroseillers,L.
TITLE New metalloproteases of the neprilysin family
JOURNAL Patent: WO 0047750-A 12 17-AUG-2000.
BOILEAU GUY (CA) ; DESGROSEILLERS LUC (CA) ; UNIVERSITE DE MONTREAL
(CA)

FEATURES
Source 1..2925
/organism="Mus musculus"
/db_xref="taxon:10090"
CDS 332..2629
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09977.1"
/db_xref="GI:10280088"
/translation="MYERAGKRRKSPGEVEGLMWLVLLLLGAIYVLGVFVSGIGQL
PLTSLSHFSWERTVVKRGLRDSLSKSDICTTPSCVIAAARILENDMSRNPENFY
QACGGLRHHRVHPETNSRYSPVDILRLDEVLKGLVEDSTQHRPAVEKAKTLFYS
CMNQSVKIDSEPLLSVLKMGWGPVAMDKNWETMGLKWELEQLAVLNSQFNRRVL
IDLFINWDDSDSRHHVYIIDQPTLGMPSREYVFOEDNNHKVKRAYLEFMTSVATMGLK
DONLSKASAMREAMEVLEETHLANATVPOEKRDHVTALVHRMDLMELOERFGLK
FWNTLEIQNVLSVEVELPDEVVYGPYLENLEDDISVSARTMQNVLYWRLVLD
RIGTSQRKEARVDTRKALIGTIVVEVRWRECSTVNSMESAVGSLIKIKRAFSLK
KTVRELIRKRSFYVDNLDELNMDESKKKAQERKADMINREIQIPDYILEDNNKHL
DEEYSSLTPEYDLFENGLQNKNAQSLIKLRKVDQNLWIIIGAAVNAFYSPPNRN
QIEYFAGILQPPFSDKQPSLNFQGGIMVYIGHEITHGFDNGRNFDKNGMLDWNNS
FSARHPQOQSCMIYQYGNSEADLNQNVNGFSTLGENIADNGGVROAYKAYLRVLA
LPGKSDORGLNLTYSQALFEINAYAQVWCGSYRPEFAVQSITKTDVHSPKRYVLGSLQ
LPGSEAFHPCRGSPGPMKRCRIW"
BASE COUNT 710. a 797 c 836 g 582 t
ORIGIN

alignment_scores:
Quality: 581.00 Length: 114
Ratio: 5.142 Gaps: 0
Percent Similarity: 99.123 Percent Identity: 85.088

alignment_block:
US-09-647-780A-4 x AX033272

Align seg 1/1 to: AX033272 from: 1 to: 2925

1 GlyHisGluIleThrHisGlyPheAspAspAsnGlyArgAsnPheAspLy 17
|||||
2123 GGGCAGAGATGACACACGCGCTTGATGATAATGCTCGTAACCTTGACAA 2172

17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
|||||
2173 GAACGGCAACATGCTGGAGTCTGAGTGAATCTTCGGCCGCGCATCTCC 2222

34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
|||||
2223 AACACAGTCGCAATGCGATGATCTATCATAGTCAGGCAACTCTCTCTGGAA 2272

51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnrl 67
|||||
2273 CTAGCAGACACCAAGATGTGAACGGATTTCATACCTCGGGAGAACAT 2322

67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
|||||
2323 TCCGACACACGAGGTGTCGCACAGGCATACAGGCTTACCTACGGTGGC 2372

84 etAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
|||||

```

```

2373 TGGCTGATGGCGGCAAGATCAGCGACTGCCGGGACTGAACCTGACCTAT 2422
101 GluGlnLeuPhePheIleAsnTyAlaGlnValTrpCysGly 114
|||||
2423 GCCCAGCTTTCTTCATCACTATGCCCAGGTGTGGTGGG 2464

seq_name: gb_ro:AF176569
seq_documentation_block:
LOCUS AF176569 2925 bp mRNA linear ROD 11-MAY-2000
DEFINITION Mus musculus neprilysin-like metalloproteinase 1 (N11) mRNA,
complete cds.
ACCESSION AF176569
VERSION AF176569.1 GI:7769082
KEYWORDS
SOURCE house mouse
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
1 (bases 1 to 2925)
AUTHORS Ghadjar,G., Ruchon,A.F., Carpentier,M., Marcinkiewicz,M.,
Seidah,N.G., Crine,P., DesGroselliers,L. and Boileau,G.
TITLE Molecular cloning and biochemical characterization of a new mouse
testis soluble zinc-metalloproteinase of the neprilysin family
JOURNAL Biochem. J. 347, 419-429 (2000)
REFERENCE 2 (bases 1 to 2925)
AUTHORS Ghadjar,G., Ruchon,A.F., DesGroselliers,L. and Boileau,G.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-1999) Biochemistry, University of Montreal, 2900
Edouard Montpetit, Montreal, Que H3T 1J4, Canada
FEATURES
source 1. 2925
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="testis"
gene 1. 2925
/genes="N11"
CDS 332..2629
/genes="N11"
/notes="N11; metallo-endopeptidase"
/codon_start=1
/product="neprilysin-like metalloproteinase 1"
/protein_id="AAF69247.1"
/db_xref="GI:7769083"
/translation="MYERAGWCRKKSFGVEYGLWLVLLLLLGAIYVTLGVYFSTGKQL
PLTSLHFSWDBETVVKRALRDSLSKSDICTTPSCVIAARILENDQSRNCFNYR
QYACGWLRRHVHPETNSRYSFVDILLRDELEVILKGVLEDSOHRPAVEKAKTLVRS
CMNQSVTEKRDSPLLSLVKMGVWPVAMDKNETMGLKWELEKALVLEFMTSVATMLRK
IDLFINWDDNSRRHVYI DQPTLGMPSREYVPEQDNHKKVKAYLEFMTSVATMLRK
DONLSKESAMVREMAEVLETHFLANATVPQEKRDHVTALYHRMDLMELQERFGLKG
FNWTLFIQNVLSSEVELFDEEVVYGIPLYLEDEI IDVSARTMYNLYLVRLVLD
RIGSLGRKEARVDYKALYGTVEEVREKCVSYVNSNESAVASGLYIKRAFESKDS
KSTVRLEIERISRVYNDLDELNMDEESKKCAEQAMNIREQIGPYDILLEONNHL
DEYVSFLTYFDFYFENGLQNNKQAKRLKREKVDQNLWIGIAAVNAFYSPNNRK
QIVFPAGILOPPFESKDQPSLGNFGIGMWIGIETHGFDNGRNFCDKNGMLDWSN
FSARHFQOOSQIMYQGNFSWELADNONGVSTGLCENTADNGGVQKAYKAYLRWL
DGKDKQLRGLNTYAGLPIFYNTAQVWCYSYRFEFAVQSIKTDVHSPCLKYRVLSGLN
LPQSFAPHCPRGSPMHMRCRIW"
BASE COUNT 710 a 797 c 836 g 582 t
ORIGIN

```

```

alignment_scores:
  Quality: 581.00      Length: 114
  Ratio: 5.142        Gaps: 0
  Percent Similarity: 99.123  Percent Identity: 85.088

alignment_block:
  US-09-647-780A-4 x AF176569

  Align seg 1/1 to: AF176569 from: 1 to: 2925

1 GLVHISGLLTLERHHSIGLVPEASPAASNGLYAFQAQSNPHEA

```

```
|||||
2123 GGGCAGGATCACACGGCTTTGATGATATATGGTCGTAACCTTTGACAA 2172
17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
|||||
2173 GAACGGCAACATGCTGGACTGCTGGAGTAACCTCTCGGCCCGGCACCTCC 2222
34 rGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
|||||
2223 AACACAGTCGCAATGCATGATCTATCAGTACGGCAACTCTCTTGGGAA 2272
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnI 67
|||||
2273 CTAGCAGACAACCAAGATGTGACGGATTTCAGTACCTCGGGGAGACAT 2322
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
|||||
2323 TGCCGACAACGGAGGTGTCCGACAGGCATACAAAGGCTTACCTACGGTGGC 2372
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
|||||
2373 TGGCTGATGGCGGCAAGATCAGCGACTCCGGGACTGAACCTGACCTAT 2422
101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114
|||||
2423 GCCCAGCTTTTCTTCATCAACTATGCCCAGGTGTGGTGGG 2464
```

OM of: US-09-647-780A-4 to: N_Geneseq_032802.* out_format : pfs

Date: Oct 6, 2002 3:43 AM

About: Results were produced by the GenCore software, version 4.5.

Copyright: (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=framet_p2n.model -DEV=xlh  
-O/cg2_1/USPTO.spool/US05647780/runat_03102002_110427_24643/app_query.fasta_1.175  
-DB=N_Geneseq_032802 -QPMT=fastcap -SUFFIX=p2n.rng -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -XGAPOP=10.000 -XGAPEXT=0.500  
-XGAPOP=4.500 -XGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum2  
-TRANS=human40 cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct  
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US05647780.@CGI_1.0 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1
```

Search information block:

Query: US-09-647-780A-4

Query length: 116

Database: N_Geneseq_032802.*

Database sequences: 1736436

Database length: 858457221

Search time (sec): 212.780000

score_list:

Sequence	Strd Orig	zScore	EScore	Len	Documentation
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000A.DAT:AAF89737 +		651.00	1443.99	4.4e-72	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF89737 +		651.00	1443.99	4.4e-72	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF89738 +		651.00	1442.80	5.1e-72	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF89660 +		651.00	1441.63	6.0e-72	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF89661 +		651.00	1441.53	6.0e-72	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000A.DAT:AAF89661 +		651.00	1441.48	6.1e-72	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF89659 +		651.00	1441.34	6.2e-72	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002A.DAT:AAF897186 +		640.50	1419.53	1.0e-70	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF896248 +		603.50	1342.65	1.9e-66	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF895007 +		588.50	1307.87	1.7e-64	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000A.DAT:AAF8963763 +		581.00	1282.32	4.5e-63	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999A.DAT:AAF898810 +		578.00	1276.09	9.9e-63	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988A.DAT:AAF8981690 +		471.00	1032.19	3.8e-49	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988A.DAT:AAF890952 +		469.00	1028.24	6.3e-49	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994A.DAT:AAF8962678 +		469.00	1028.24	6.3e-49	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF894514 +		469.00	1028.24	6.3e-49	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF894524 +		469.00	1028.24	6.3e-49	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF894073 +		469.00	1028.24	6.3e-49	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF890123 +		467.00	1018.26	2.3e-48	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992A.DAT:AAF892244 +		465.00	1032.57	3.6e-49	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988A.DAT:AAF892261 +		465.00	1019.00	2.1e-48	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994A.DAT:AAF892679 +		465.00	1019.00	2.1e-48	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAF8910163 +		397.50	871.22	3.5e-40	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:AAF8910162 +		378.50	820.53	2.3e-37	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995A.DAT:AAF893868 +		374.50	831.56	5.7e-38	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995A.DAT:AAF893874 +		374.50	818.53	3.0e-37	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995A.DAT:AAF893860 +		374.50	817.71	3.4e-37	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998A.DAT:AAF896149 +		370.50	806.47	1.4e-36	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996A.DAT:AAF8942151 +		367.50	801.35	2.7e-36	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996A.DAT:AAF893199 +		367.50	801.35	2.7e-36	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995A.DAT:AAF893877 +		367.50	800.98	2.9e-36	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995A.DAT:AAF893811 +		367.50	800.28	3.1e-36	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995A.DAT:AAF893861 +		367.50	800.28	3.1e-36	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998A.DAT:AAF8923264 +		353.00	766.10	2.5e-34	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000A.DAT:AAF8912670 +		353.00	766.10	2.5e-34	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999A.DAT:AAF8934329 +		343.50	752.72	1.4e-33	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000A.DAT:AAF898595 +		343.50	752.72	1.4e-33	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF890579 +		343.50	750.86	1.8e-33	
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF890422 +		343.50	746.26	3.2e-33	

/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000A.DAT:AAF89402 + 343.50 745.04 3.8e-33
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999A.DAT:AAF894326 + 339.50 737.41 1.0e-32
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000A.DAT:AAF89592 + 339.50 737.41 1.0e-32
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF896134 + 339.50 737.41 1.0e-32
/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998A.DAT:AAF891976 + 339.50 735.08 1.3e-32

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF89737

seq_documentation_block:

ID AAF89737 standard; DNA; 2076 BP.

XX

AC AAF89737;

XX

DT 23-JUL-2001 (first entry)

XX

DE Nucleotide sequence of a human metalloprotease enzyme IGS5.

XX

Metalloprotease; IGS5; infection; pain; cancer; diabetes; obesity;
anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;
hypotension; hypertension; urinary retention; osteoporosis;
angina pectoris; myocardial infarction; stroke; ulcer; allergy;
benign prostatic hypertrophy; migraine; psychotic disorder;
neurological disorder; autism; multiple sclerosis; Alzheimer's disease;
neurodegenerative disease; sleep disorder; epilepsy; kidney disease;
cardiovascular disease; arteriosclerosis; cerebrovasospasm;
subarachnoid hemorrhage; cerebral ischemia; cerebral infarction;
peripheral vascular disease; Raynaud's disease; motility disorder;
gastrointestinal disorder; gastric emptying; gastroparesis; diarrhoea;
inflammation; chemotherapy induced injury; tumour invasion;
immune disorder; arthritis; endotoxin shock; sepsis; diabetes mellitus;
severe mental retardation; dyskinesia; Huntington's disease;
Gilles de la Tourette's syndrome; ss.

XX

OS Homo sapiens.

XX

Key Location/Qualifiers

XX CDS

1..2076

/*tag= a

/product= "metalloprotease enzyme IGS5"

XX WO200136610-A1.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-EPI1532.

XX 19-NOV-1999; 99EP-0203862.

XX 19-NOV-1999; 99NL-1013616.

XX 31-MAY-2000; 2000EP-0201937.

XX 31-MAY-2000; 2000NL-1015356.

XX (SOLV) SOLVAY PHARM BV.

XX Deleersnijder W, Wiegiers R, Weske M;

XX WPI; 2001-343815/36.

XX P-PSDB; AAB83840.

XX New IGS5 polypeptides useful for treating infections, pain, cancer,

XX diabetes, obesity, anorexia, bulimia, asthma, schizophrenia,

XX hypertension, urinary retention and Parkinson's disease

XX Claim 11; Page 5-6; 115pp; English.

XX The present sequence encodes a human metalloprotease enzyme designated

XX IGS5. IGS5 polynucleotides and polypeptides are useful for treating

XX infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma,

XX Parkinson's disease, acute heart failure, hypotension, hypertension,

XX urinary retention, osteoporosis, angina pectoris, myocardial infarction,

XX stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,

XX psychotic and neurological disorders, autism, multiple sclerosis,

XX Alzheimer's disease, and other neurodegenerative diseases, sleep

XX disorder, epilepsy, cardiovascular diseases, arteriosclerosis,

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSertTrpasp 50
|||||
1856 GGGAGCAGTCAGAGTCATGATCTACAGTACGCAACTACTCTGGGAC 1905

51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsn11 67
|||||
1906 CTGGCAGACGACAGACAGCTGACGAGGATCAACACCCCTGGGGAAACAT 1955

67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
|||||
1956 TGCTGACACGAGGGGTGGCGCAAGCCCTATAGGCTTACCTCAAGTGA 2005

84 eTalaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
|||||
2006 TGGCAGAGGGTGGCAAGGACGACAGCTGCCCGCTGGATCTCACCCAT 2055

101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114
|||||
2056 GAGCAGCTCTTCTTCATCACTACGACGAGGTGGTGGCGG 2097

seq_name: /SDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF89738

seq_documentation_block:

ID AAF89738 standard; DNA: 2340 BP.

XX AC AAF89738;

XX DT 23-JUL-2001 (first entry)

XX DE Nucleotide sequence of a human metalloprotease enzyme IG55.

XX KW Metalloprotease; IG55; infection; pain; cancer; diabetes; obesity;
KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;
KW hypertension; hypertension; urinary retention; osteoporosis;
KW angina pectoris; myocardial infarction; stroke; ulcer; allergy;
KW benign prostatic hyperplasia; migraine; psychotic disorder;
KW neurological disorder; autism; multiple sclerosis; Alzheimer's disease;
KW neurodegenerative disease; sleep disorder; epilepsy; kidney disease;
KW cardiovascular disease; arteriosclerosis; cerebrovasospasm;
KW subarachnoid hemorrhage; cerebral ischemia; cerebral infarction;
KW peripheral vascular disease; Raynaud's disease; motility disorder;
KW gastrointestinal disorder; gastric emptying; gastroparesis; diarrhoea;
KW inflammation; chemotherapy induced injury; tumour invasion;
KW immune disorder; arthritis; endotoxin shock; sepsis; diabetes mellitus;
KW severe mental retardation; dyskinesia; Huntington's disease;
KW Gilles de la Tourette's syndrome; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 1..2340

XX FT /*tag= a

XX FT /product= "metalloprotease enzyme IG55"

XX PN WO200136610-A1.

XX PD 25-MAY-2001.

XX PF 17-NOV-2000; 2000WO-EP11532.

XX PR 19-NOV-1999; 99EP-0203862.

XX PR 19-NOV-1999; 99NL-1013616.

XX PR 31-MAY-2000; 2000EP-0201937.

XX PR 31-MAY-2000; 2000NL-1015356.

XX PA (SOLV) SOLVAY PHARM BV.

XX PI Deleersnijder W, Wiegiers R, Weske M;

XX DR WPI; 2001-343815/36.

XX DR P-PSDB; AAB83841.

XX DT

PT New IG55 polypeptides useful for treating infections, pain, cancer,
PT diabetes, obesity, anorexia, bulimia, asthma, schizophrenia;
PT hypertension, urinary retention and Parkinson's disease
XX
XX Claim 11; Page 6-7; 115pp; English.
XX
CC The present sequence encodes a human metalloprotease enzyme designated
CC IG55. IG55 polynucleotides and polypeptides are useful for treating
CC infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma,
CC Parkinson's disease, acute heart failure, hypotension, hypertension,
CC urinary retention, osteoporosis, angina pectoris, myocardial infarction,
CC stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,
CC psychotic and neurological disorders, autism, multiple sclerosis,
CC Alzheimer's disease, and other neurodegenerative diseases, sleep
CC disorder, epilepsy, cardiovascular diseases, arteriosclerosis,
CC cerebrovasospasm, subarachnoid hemorrhage, cerebral ischemia, cerebral
CC infarction, peripheral vascular disease, Raynaud's disease, kidney
CC diseases, gastrointestinal disorders, motility disorders and conditions
CC of delayed gastric emptying, post-operative or diabetic gastroparesis,
CC diarrhoea, inflammations, chemotherapy induced injury, tumour invasion,
CC immune disorders, arthritis, endotoxin shock, sepsis, complications of
CC diabetes mellitus, and severe mental retardation and dyskinesias, such
CC as Huntington's disease or Gilles de la Tourette's syndrome.

XX SQ Sequence 2340 BP; 539 A; 649 C; 739 G; 413 T; 0 other;

alignment_scores:
Quality: 651.00 Length: 114
Ratio: 5.711 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-647-780A-4 x AAF89738

Align seg 1/1 to: AAF89738 from: 1 to: 2340

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAspLy 17
|||||
1834 GGGCAGAGATCAGCAGCGCTTTGACGACANTGGCGGAACTTCGACAA 1883

17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
|||||
1884 GAATGGCAACATGATGATGGTGGAGTAACTTCTCCACCCAGCACATCC 1933

34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSertTrpasp 50
|||||
1934 GGGAGCAGTCAGATGATCTACAGTACGCGCAACTACTCTCTGGGAC 1983

51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsn11 67
|||||
1984 CTGGCAGACGACAGACAGCTGACGAGGATCAACACCCCTGGGGAAACAT 2033

67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
|||||
2034 TGCTGACACGAGGGGTGGCGCAAGCCCTATAGGCTTACCTCAAGTGA 2083

84 eTalaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
|||||
2084 TGGCAGAGGGTGGCAAGGACGACAGCTGCCCGCTGGATCTCACCCAT 2133

101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114
|||||
2134 GAGCAGCTCTTCTTCATCACTACGACGAGGTGGTGGCGG 2175

seq_name: /SDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAF59660

seq_documentation_block:

ID AAF59660 standard; cDNA: 2636 BP.

XX AC AAF59660;

XX DT 27-APR-2001 (first entry)

xx DE Human neprilysin-like membrane metalloproteinase SNEPB cDNA.

xx KW Human; SNEPB; neprilysin-like membrane metalloproteinase;

xx KW splice variant; alternative splicing; zinc endopeptidase family;

xx KW neuropeptide; peptide hormone; processing; metabolism; vaccine;

xx KW drug screening; renal insufficiency; hepatic disease; ischaemia; pain;

xx KW hypertension; cancer; inflammation; cardiovascular disease;

xx KW neuronal disease; pancreatic disease; prostatic disease;

xx KW respiratory disease; gene therapy; nephrotropic; vasotropic; analgesic;

xx KW cerebroprotective; hypotensive; cytostatic; antiinflammatory;

xx KW cardiovascular; hepatotropic; ss.

OS Homo sapiens.

xx KW EPI069188-AL.

xx PD 17-JAN-2001.

xx KW 15-JUL-1999; 99EP-0401767.

xx KW 15-JUL-1999; 99EP-0401767.

xx KW (SNFI) SANOFI-SYNTHELABO.

xx KW Jagerschmidt A, Agnel M, Culouscou J;

xx KW WPI: 2001-212582/22.

xx KW P-PSDB; AAB60562.

xx KW New membrane-associated metalloproteinase SNEPA, SNEPB and SNEPC

xx KW polypeptides and polynucleotides, useful for treating e.g. acute and

xx KW chronic renal insufficiency, pain, stroke, cancer, inflammation, renal

xx KW and hepatic ischemia

PS Claim 5; Page 30-33; 72pp; English.

xx KW The invention relates to the human SNEPA, SNEPB and SNEPC proteins, and

xx KW the cDNAs encoding them. SNEPA, SNEPB and SNEPC are neprilysin-like

xx KW membrane metalloproteinases and are the products of alternative splicing.

xx KW The substrate(s) for the SNEP proteins are not as yet known, although

xx KW the neprilysin family of zinc endopeptidases play key roles in the

xx KW processing and/or metabolism of neuropeptides and peptide hormones. SNEP

xx KW nucleotides may be used as hybridisation probes for cDNA and genomic

xx KW DNA; to isolate full-length cDNAs and genomic clones encoding SNEPA,

xx KW SNEPB or SNEPC; to isolate cDNA and genomic clones of SNEP homologues;

xx KW as research reagents and material for the discovery of treatments and

xx KW diagnostics for animal and human diseases; and for chromosome

xx KW identification. The SNEP proteins may be used as immunogens to

xx KW produce antibodies immunospecific for SNEPA, SNEPB or SNEPC. Such

xx KW antibodies are used to isolate or identify clones expressing the

xx KW protein, or to purify the proteins by affinity chromatography.

xx KW SNEP proteins may also be used in screening for compounds which modulate

xx KW SNEP endopeptidase activity, and to assess enzymatic cleavage of small

xx KW molecule substrates in cells, cell-free preparations, chemical libraries

xx KW and product mixtures. The SNEP proteins (as vaccine compositions),

xx KW SNEP nucleotides, and SNEP activators or inhibitors may be used

xx KW to treat acute and chronic renal insufficiency, cancer, inflammation, as

xx KW ischaemia, pain, stroke, hypertensive disease, renal, inflammatory, as

xx KW well as cardiovascular, neuronal, pancreatic, prostatic, renal,

xx KW respiratory or hepatic diseases. They may also be used in modulating

xx KW peptide activation and/or degradation in the brain or kidney or in

xx KW another organ, or to diagnose or treat any disorder related to abnormal

xx KW expression of SNEPA, SNEPB or SNEPC. The present sequence represents

xx KW cDNA encoding SNEPB.

SQ Sequence 2636 BP; 601 A; 759 C; 813 G; 462 T; 1 other;

alignment_scores:

Quality:	651.00	Length:	114
Ratio:	5.711	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-647-780A-4 x AAF59660.

Align seg 1/1 to: AAF59660 from: 1 to: 2636

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAspLy 17

1763 GGGCAGCAGATCAGCAGCGCTTTGACGACAAATGGCGGAATTCGACAA 1812

17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34

1813 GAATGCAACATGATGGATTGGTGGAGTAACCTTCTCCACCCAGCAGCTTC 1862

34 rGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50

1863 GGGAGCAGTCAGAGTCATGATCTACCACTAGCGCAACTACTCTCTGGGAC 1912

51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnI 67

1913 CTGGCAGACGACAGACAGCTGACGAGGATTCACACCCCTTGGGGAACAT 1962

67 eAlaAspAsnGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84

1963 TGTGACAAACGAGGGGTGGCGCAAGCCTATAAGGCTACTCTCAAGTGA 2012

84 eTalaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100

2013 TGGCAGAGGTGGCAAGGACGACGAGCTGCGCGGCTGGATCTCACCCT 2062

101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114

2063 GAGCAGCTCTTCTTCATCACTATGCCAGGTGCTGGTGGCGG 2104

seq_name: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT-AAF59661

seq_documentation_block:

ID AAF59661 standard; cDNA; 2663 BP.

XX AAF59661;

XX 27-APR-2001 (first entry)

XX Human neprilysin-like membrane metalloproteinase SNEPB cDNA.

XX Human; SNEPB; neprilysin-like membrane metalloproteinase;

XX KW splice variant; alternative splicing; zinc endopeptidase family;

XX KW neuropeptide; peptide hormone; processing; metabolism; vaccine;

XX KW drug screening; renal insufficiency; hepatic disease; ischaemia; pain;

XX KW hypertension; cancer; inflammation; cardiovascular disease;

XX KW neuronal disease; pancreatic disease; prostatic disease;

XX KW respiratory disease; gene therapy; nephrotropic; vasotropic; analgesic;

XX KW cerebroprotective; hypotensive; cytostatic; antiinflammatory;

XX KW cardiovascular; hepatotropic; ss.

OS Homo sapiens.

XX KW EPI069188-AL.

XX PD 17-JAN-2001.

XX KW 15-JUL-1999; 99EP-0401767.

XX KW 15-JUL-1999; 99EP-0401767.

XX KW (SNFI) SANOFI-SYNTHELABO.

XX KW Jagerschmidt A, Agnel M, Culouscou J;

XX KW WPI: 2001-212582/22.

XX KW P-PSDB; AAB60563.

XX KW New membrane-associated metalloproteinase SNEPA, SNEPB and SNEPC

PT polypeptides and polynucleotides, useful for treating e.g. acute and
PT chronic renal insufficiency, pain, stroke, cancer, inflammation, renal
PT and hepatic ischemia.

XX Claim 5; Page 35-38; 72pp; English.

XX The invention relates to the human SNEPb, SNEPb and SNEPc proteins, and
CC the cDNAs encoding them. SNEPb, SNEPb and SNEPc are neprilysin-like
CC membrane metalloproteinases and are the products of alternative splicing.
CC The substrate(s) for the SNEP proteins are not as yet known, although
CC the neprilysin family of zinc endopeptidases play key roles in the
CC processing and/or metabolism of neuropeptides and peptide hormones. SNEP
CC nucleotides may be used as hybridisation probes for cDNA and genomic
CC DNA; to isolate full-length cDNAs and genomic clones encoding SNEPb,
CC SNEPb or SNEPc; to isolate cDNA and genomic clones of SNEP homologues;
CC as research reagents and material for the discovery of treatments and
CC diagnostics for animal and human diseases; and for chromosome
CC identification. The SNEP proteins may be used as immunogens to
CC produce antibodies immunospecific for SNEPb, SNEPb or SNEPc. Such
CC antibodies are used to isolate or identify clones expressing the
CC protein, or to purify the proteins by affinity chromatography.
CC SNEP proteins may also be used in screening for compounds which modulate
CC SNEP endopeptidase activity, and to assess enzymatic cleavage of small
CC molecule substrates in cells, cell-free preparations, chemical libraries
CC and product mixtures. The SNEP proteins (as vaccine compositions),
CC SNEP nucleotides, and SNEP activators or inhibitors may be used
CC to treat acute and chronic renal insufficiency, renal and hepatic
CC ischaemia, pain, stroke, hypertensive disease, cancer, inflammation, as
CC well as cardiovascular, neuronal, pancreatic, prostatic, renal,
CC respiratory or hepatic diseases; they may also be used in modulating
CC peptide activation and/or degradation in the brain or kidney or in
CC another organ, or to diagnose or treat any disorder related to abnormal
CC expression of SNEPb, SNEPb or SNEPc. The present sequence represents
CC cDNA encoding SNEPc.

XX Sequence 2663 BP; 609 A; 769 C; 816 G; 468 T; 1 other:

alignment_scores: Quality: 651.00 Length: 114
Ratio: 5.711 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-647-780A-4 x AAF59661

Align seg 1/1 to: AAF59661 from: 1 to: 2663

1 GlyHisGluIleThrHisGlyPheAspAspAsnGlyArgAsnPheAspLY 17
1790 GGGCAGCAGATCATCGGCGCTTTGACGACATGCGCGGACATTCGACAA 1839

17 sasnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
1840 GAATGCAACATGATGATGTTGGTGGAGTAACTTCTCCACCCAGCACTTC 1889

34 rGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
1890 GGGCAGGAGTCAGAGTCAGTATCTACAGTACGCACTACTCTCTGGGAC 1939

51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnII 67
1940 CTGGCAGACGACAGACGATGACGAGGATTCACACCCCTTGGGGAACAT 1989

67 eAlaAspAsnGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
1990 TGGTGAACACGAGGGGTGCGGCAACCTTAAAGGCTTACCTCAAGTGA 2039

84 eAlaGluGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
2040 TGGCAGAGGGTGGCAGGACCAAGCAGCTGCCGCGCTGGATCTCAACCAT 2089

101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly. 114

|||||
2090 GAGCAGCTCTCTTCATCACTATGCCAGGTGTGGTGGG 2131

seq_name: /SDSI/gcdata/geneseq/geneseq-emb1/NA2000.DAT:AAA63764

seq_documentation_block:

ID AAA63764 standard; cDNA; 2676 BP.

XX AAA63764;

DT 04-DEC-2000 (first entry)

XX cDNA encoding neutral endopeptidase metalloproteinase-like enzyme NL-2.

XX Neprilysin; neutral endopeptidase metalloproteinase-like enzyme;

KW NEP-like enzyme; protein production; protein secretion;

KW neurological disease; Alzheimer's disease; pain; psychiatric disorder;

XX fertility; bone disease; abnormal phosphate metabolism; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 7..2319

FT /*tag= a

FT /product= "neutral endopeptidase metalloproteinase-like

FT enzyme NL-2"

XX WO200047750-A2.

XX 17-AUG-2000.

XX 11-FEB-2000; 2000WO-CA00147.

XX 11-FEB-1999; 99CA-2260376.

XX (UYMO-) UNIV MONTREAL.

XX Desgroseillers L, Boileau G;

DR WPI; 2000-549148/50.

DR P-PSDB; AAB08131.

XX Novel neutral endopeptidase-like metalloproteinase polypeptides and

PT polynucleotides, used to screen for related sequences and enzyme

PT inhibitors, used for the treatment of NL-3 related bone disorders

XX Disclosure: Fig 4; 59pp; English.

XX The present sequence encodes a human neutral endopeptidase

CC metalloproteinase-like enzyme, designated NL-2. The specification

CC also describes NL-1 and NL-3. The NL enzymes are used to test for

CC specific inhibitors. The N-terminal region of the enzymes can be used

CC to promote production and secretion of foreign proteins and active

CC biopeptides, using chimeric constructs containing the foreign protein

CC downstream from and in phase with the N-terminal region. The NL enzymes

CC have been localised to the brain, and may be useful in the

CC treatment of neurological diseases such as Alzheimer's disease, pain,

CC and psychiatric disorders. NL enzymes have also been localised to the

CC testis and ovaries, and may be used to control fertility. They have

CC also been localised to bones, and may be used to treat bone diseases,

CC and abnormal phosphate metabolisms related to improper peptide

CC processing by the NL-3 enzyme.

XX Sequence 2676 BP; 608 A; 771 C; 823 G; 474 T; 0 other:

alignment_scores:

Quality: 651.00 Length: 114

Ratio: 5.711 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-647-780A-4 x AAA63764

Align seg 1/1 to: AAA63764 from: 1 to: 2676

```
1 GlyHisGluLeuThrHisGlyPheAspAspAsnGlyArgAsnPheAspLy 17
|||||
1813 GGGCAGAGATCAGCAGCGCTTTGACGACAATGGCGGAACCTTCGACAA 1862
|||||
17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
|||||
1863 GAATGCCAACATGATGGATTGGTGGAGTAACTTCTCCACCAGCACTTCC 1912
|||||
34 rGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
|||||
1913 GGGAGCAGTCAGAGTCATGATCTACCAAGTACGCAACTACTCTCTGGAC 1962
|||||
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnII 67
|||||
1963 CTGGCAGACGACAGACAGCTGACGAGGATTCACACACCTTTGGGGAACAT 2012
|||||
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
|||||
2013 TGCTGACACAGGAGGGGTGGCGCAACCTATAGCGCTACCTCAATGGA 2062
|||||
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
|||||
2063 TGGCAGAGGTGGCAAGGACAGCAGCTGCCGCGCTGGATCTCACCCAT 2112
|||||
101 GluGlnLeuPhePheLeuAsnTyrAlaGlnValTrpCysGly 114
|||||
2113 GAGCAGCTCTTCTTCATCAACTATGCCAGGTGGTGGCGG 2154
|||||
```

seq_name: /SID1/cgdata/geneseq/geneseq-emb1/NA2001A.DAT.AAF59659

seq_documentation_block:

ID AAF59659 standard; cDNA; 2714 BP.

XX AAF59659;

27-APR-2001 (first entry)

XX Human neprilysin-like membrane metalloproteinase SNEPA cDNA.

DE Human; SNEPA; neprilysin-like membrane metalloproteinase;
KW splice variant; alternative splicing; zinc endopeptidase family;
KW neurotensin; peptide hormone; processing; metabolism; vaccine;
KW drug screening; renal insufficiency; hepatic disease; ischaemia; pain;
KW hypertension; cancer; inflammation; cardiovascular disease;
KW neuronal disease; pancreatic disease; prostatic disease;
KW respiratory disease; gene therapy; nephrotropic; vasotropic; analgesic;
KW cerebroprotective; hypotensive; cytostatic; antiinflammatory;
KW cardiovascular; hepatotropic; ss.

XX Homo sapiens.

XX EP1069188-A1.

XX 17-JAN-2001.

XX 15-JUL-1999; 99EP-0401767.

XX 15-JUL-1999; 99EP-0401767.

XX (SNFI) SANOFI-SYNTHELABO.

XX Jagerschmidt A, Agnel M, Culouscou J;

XX WPI; 2001-212582/22.

XX P-PSDB; AAB60561.

XX New membrane-associated metalloproteinase SNEPA, SNEPB and SNEPC
PT polypeptides and polynucleotides, useful for treating e.g. acute and
PT chronic renal insufficiency, pain, stroke, cancer, inflammation, renal
PT and hepatic ischemia

XX

PS Claim 5; Page 25-28; 72pp; English.

XX The invention relates to the human SNEPA, SNEPB and SNEPC proteins, and
CC the cDNAs encoding them. SNEPA, SNEPB and SNEPC are neprilysin-like
CC membrane metalloproteinases and are the products of alternative splicing.
CC The substrate(s) for the SNEP proteins are not as yet known, although
CC the neprilysin family of zinc endopeptidases play key roles in the
CC processing and/or metabolism of neuropeptides and peptide hormones. SNEP
CC nucleotides may be used as hybridisation probes for cDNA and genomic
CC DNA; to isolate full-length cDNAs and genomic clones encoding SNEPA,
CC SNEPB or SNEPC; to isolate cDNA and genomic clones of SNEP homologues;
CC as research reagents and material for the discovery of treatments and
CC diagnostics for animal and human diseases; and for chromosome
CC identification. The SNEP proteins may be used as immunogens to
CC produce antibodies immunospecific for SNEPA, SNEPB or SNEPC. Such
CC antibodies are used to isolate or identify clones expressing the
CC protein, or to purify the proteins by affinity chromatography.
CC SNEP proteins may also be used in screening for compounds which modulate
CC SNEP endopeptidase activity, and to assess enzymatic cleavage of small
CC molecule substrates in cells, cell-free preparations, chemical libraries
CC and product mixtures. The SNEP proteins (as vaccine compositions),
CC SNEP nucleotides, and SNEP activators or inhibitors may be used
CC to treat acute and chronic renal insufficiency, cancer, inflammation, as
CC ischaemia, pain, stroke, hypertensive disease, pancreatic, renal,
CC well as cardiovascular, neuronal, prostatic, renal,
CC respiratory or hepatic diseases. They may also be used in modulating
CC peptide activation and/or degradation in the brain or kidney or in
CC another organ, or to diagnose or treat any disorder related to abnormal
CC expression of SNEPA, SNEPB or SNEPC. The present sequence represents
CC cDNA encoding SNEPA.

XX SQ Sequence 2714 BP; 621 A; 780 C; 836 G; 476 T; 1 other;

alignment_scores:

Quality: 651.00 Length: 114
Ratio: 5.711 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-647-780A-4 x AAF59659

Align seg 1/1 to: AAF59659 from: 1 to: 2714

```
1 GlyHisGluLeuThrHisGlyPheAspAspAsnGlyArgAsnPheAspLy 17
|||||
1841 GGGCAGAGATCAGCAGCGCTTTGACGACAATGGCGGAACCTTCGACAA 1890
|||||
17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
|||||
1891 GAATGCCAACATGATGGATTGGTGGAGTAACTTCTCCACCAGCACTTCC 1940
|||||
34 rGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
|||||
1941 GGGAGCAGTCAGAGTCATGATCTACCAAGTACGCAACTACTCTCTGGAC 1990
|||||
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnII 67
|||||
1991 CTGGCAGACGACAGACAGCTGACGAGGATTCACACACCTTTGGGGAACAT 2040
|||||
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
|||||
2041 TGCTGACACAGGAGGGGTGGCGCAACCTATAGCGCTACCTCAATGGA 2090
|||||
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
|||||
2091 TGGCAGAGGTGGCAAGGACAGCAGCTGCCGCGCTGGATCTCACCCAT 2140
|||||
101 GluGlnLeuPhePheLeuAsnTyrAlaGlnValTrpCysGly 114
|||||
2141 GAGCAGCTCTTCTTCATCAACTATGCCAGGTGGTGGCGG 2182
|||||
```

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:AAS97186

seq_documentation_block:

ID AAS97186 standard; cDNA; 2232 BP.

XX AC AAS97186;

XX 26-FEB-2002 (first entry)

XX Human metalloprotease partial DNA sequence #15.

XX Human; protease; PCR primer; cytostatic; immunomodulator; cardiant;
KW vasotrophic; antimigraine; analgesic; endocrine; nootropic; tranquiliser;
KW hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic;
KW anorectic; antiinflammatory; aspartyl protease; cysteine protease;
KW metalloprotease; serine protease; cancer; haematopoietic; breast; colon;
KW lung; prostrate; cervical; brain; ovarian; bladder; kidney; pain;
KW immune-related disease; cardiovascular disease; neuronal disease;
KW migraine; sexual dysfunction; mood disorder; attention disorder;
KW cognition disorder; hypotension; hypertension; psychotic disorder;
KW dyskinesia; metabolic disorder; inflammatory disorder; ss.

XX Homo sapiens.

XX WO200183782-A2.

XX 08-NOV-2001.

XX 04-MAY-2001; 2001WO-US14431.

XX 04-MAY-2000; 2000US-201879P.

XX (SUGE-) SUGEN INC.

XX Plowman CD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;

XX Payne V;

XX WPI; 2002-041502/05.

XX P-PSDB; AAU72903.

XX Novel protease polypeptide useful for screening for substances that may
PT be used to treat, e.g., cancers, immune-related diseases,
PT cardiovascular disease, migraine, pain, psychotic and inflammatory
PT disorders

XX Claim 30; Figure 1R-S; 232pp; English.

XX The invention relates to an isolated, enriched, or purified protease
CC polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to
CC screen for substances (S) that may modulate its activity. Administering
CC S (which modulates protease activity in vitro) may be used to treat a
CC disease or disorder selected from cancers (e.g., of tissues, of blood or
CC haematopoietic origin, of the breast, colon, lung, prostate, cervical,
CC brain, ovarian, bladder or kidney), immune-related diseases and
CC disorders, cardiovascular disease, brain or neuronal-associated diseases
CC (e.g., central or peripheral nervous system diseases, migraine, pain,
CC sexual dysfunction, mood disorders, attention disorders, cognition
CC disorders, hypotension, hypertension, psychotic disorders, neurological
CC disorders and dyskinesias), metabolic disorders and inflammatory
CC disorders. (I) may also be useful as a diagnostic tool for a disease or
CC disorder such as those above. AAS97159-AAS97195 represent human
CC protease coding sequences and primers of the invention.

XX Sequence 2232 BP; 512 A; 620 C; 705 G; 395 T; 0 other;

alignment_scores:

Quality: 640.50 Length: 115
Ratio: 5.618 Gaps: 1
Percent Similarity: 99.130 Percent Identity: 99.130

alignment_block:

US-09-647-780A-4 x AAS97186

Align seg 1/1 to: AAS97186 from: 1 to: 2232

1 GlyHisGluIleThrHisGlyPheAspAspAsn...GlyArgAsnPheAs 16
1723 GGGCAGGAGATCATCGCACGCGCTTTGACGACAATGGTGGCGGAATTCGA 1772
16 PysAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisP 33
1773 CAAGAATGCCAACATGATGGATTGGTGGAGTAACCTTCTCCACCAGCACT 1822
33 heArgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTip 49
1823 TCGGGAGCAGTCAGAGTCATGATCTACCACTACGCACTACTCTCTGG 1872
50 AspLeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAs 66
1873 GACCTGGCAGACCAACAGAACGTCGAACGGATTCAACACCCCTTGGGGA 1922
66 nileAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLys 83
1923 CATGTCTGACACGAGGGGTGGCGCAAGCCTATAAGGCGCTCACTCAAGT 1972
83 rpMetAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThr 99
1973 GGATGGCAGAGGGTGGCAAGGACCAAGCAGCTGCCCGCGCTGGATCTCACC 2022
100 HisGluGlnLeuPhePheLeuAsnTyrAlaGlnValTrpCysGly 114
2023 CATGACGAGCTCTCTTCATCAACTATGCCAGGTGTGGTGGCGGG 2067
seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH26248

seq_documentation_block:

ID AAH26248 standard; cDNA; 1124 BP.

XX AC AAH26248;

XX 02-OCT-2001 (first entry)

XX Human secreted protein cDNA clone HSIDD62.

XX Secreted protein; human; HSIDD62; colon cancer; ovary cancer;
KW tumour; diagnosis; gene therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FT CDS 2..367
FT /tag= a
FT /partial

XX WO200153343-A1.

XX 26-JUL-2001.

XX 17-JAN-2001; 2001WO-US01436.

XX 18-JAN-2000; 2000US-0176307.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Shi Y;

XX WPI; 2001-483137/52.

XX P-PSDB; AAB82596.

XX Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's
PT diseases and cancers

XX Claim 1(a); Page 330; 339pp; English.

CC The present sequence is that of cDNA clone HSIDD62 (deposited as
CC ATCC PTAL201), which encodes a novel human secreted protein (see
CC AAB82596) that shows homology to atrial natriuretic protein
CC binding protein (ANBP), which is thought to have diuretic
CC (particularly natriuretic) and hypotensive activities. The gene
CC is expressed in adult small intestine tissue as well as in colon
CC tumour tissue and in ovarian cancer tissue. This tissue
CC distribution and homology to human ANBP indicate that
CC polynucleotides, translation products and antibodies corresponding
CC to this gene could be useful in the diagnosis, detection and/or
CC treatment of diseases and/or disorders of the gastrointestinal
CC system, as well as colon tumours and ovarian cancer. Antagonists
CC directed against the gene translation product may be useful in
CC prohibiting or eliminating its biological activities, such as
CC diuretic and hypotensive activities. Vectors, host cells,
CC antibodies, recombinant methods for producing the secreted protein,
CC and screening methods are also provided.
XX
SQ Sequence 1124 BP; 251 A; 362 C; 320 G; 191 T; 0 other;

alignment_scores:
Quality: 603.50 Length: 152
Ratio: 5.294 Gaps: 2
Percent Similarity: 75.000 Percent Identity: 74.342

alignment_block:
US-09-647-780A-4 x AAH26248

Align seg 1/1 to: AAH26248 from: 1 to: 1124

```
1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17
|||||
140 GGGCAGCAGATCACCGCGCTTGCACCAATGCCGGAACCTCGACAA 189

17 sAsnGlyAsnMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
|||||
190 GAATGGCAACATGATGGTGGTGGAGTAACCTCCACCCAGCACTTCC 239

34 rGluGlnSerGluCysMetIleThrGlnThrGlyAsnTrpSerTrpAsp 50
|||||
240 GGGACAGCTCAGAGTGCATGATCTACCACTAGTACGGCACTACTCTCTGGGAC 289

51 LeuAlaAspGluGln. Asn. .... 56
|||||

290 CTGGCAGACGACACAGACAGTGCCTGCCACCCAGCAGCCGCTCGGG 339

56 ..... 56

340 GTACCGGAGCCCGCCAGCTGCGCTGAGGGAGAGGGAAGTCAGGSCCG 389

57 ..... ValAsnGlyPheAsnThrLe 63
|||||

390 GGCTCGCCCAATFCCTCTCTCTGCGCAGGTGTAACGGATTCAACACCT 439

63 uGlyGluAsnIleAlaAspAsnGlyGlyValArgGlnAlaTrpLysAla 80
|||||
440 TGGGAARACATTCGTGACACGAGGGGTGCGCAANGCCTATAGGCCT 489

80 TrpLeuTrpMetAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeu 96
|||||
490 ACCTCAAGTGGATGGCAGAGGTGGCAAGGACCAGCAGCTGCCGCCGCTG 539

97 AspLeuThrHisGluGlnLeuPhePheIleAsnTrpAlaGlnValTrpCy 113
|||||
540 GATCTCACCCATGAGCAGCTCTCTCTCATCACTATGCCAGGTGTGGTG 589

113 sGly 114
||||
590 CGGG 593
```

seq_name: /SID1/gcgdata/geneseq/geneseq - emb1/NA2001A.DAT:AAH35007

seq_documentation_block:

ID AAH35007 standard; cDNA; 1226 BP.

AC AAH35007;

XX 03-SEP-2001 (first entry)

XX Human colon cancer antigen encoding cDNA SEQ ID NO:2089.

XX Human colon cancer antigen; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma; ss.

OS Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.

XX 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX P-PSDB; AAG75602.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers.

XX Claim 1; Page 3574; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where
XX the proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytostatic activity and can be used in gene
XX therapy and vaccine production. N and P may be used in the prevention,
XX diagnosis and treatment of diseases associated with inappropriate P
XX expression. For example, N and P may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of P by expressing P.
XX Inactive proteins or to supplement the patients own production of P.
XX Additionally, N may be used to produce the colon cancer-associated Ps,
XX by inserting the nucleic acids into a host cell and culturing the cell
XX to express the proteins. N and P can be used in the prevention, diagnosis
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX and AAB77789 represent sequences used in the exemplification of the
XX present invention.

XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were
XX missing at time of publication, meaning no sequences are present for
XX SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 1226 BP; 270 A; 388 C; 345 G; 213 T; 10 other;

alignment_scores:

Quality: 588.50 Length: 151
Ratio: 5.254 Gaps: 1
Percent Similarity: 74.172 Percent Identity: 74.172

alignment_block:

US-09-647-780A-4 x AAH35007

Align seg 1/1 to: AAH35007 from: 1 to: 1226

```
1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17
|||||
140 GGGCAGCAGATCACCGCGCTTTCACCAATGCCGGAACCTCGACAA 189

17 sAsnGlyAsnMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
```

```

|||||
190 GAATGGCAACATGATGGATTGGGAGTAACTTCTCCACCCAGCACTTCC 239
|||||
34' rGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrrpasp 50
|||||
240 GGGAGCAGTCAGAGTCGATGATACCACTACGAGGCAACTACTCTCTGGGAC 289
|||||
51 LeuAlaaspGluGlnasn..... 56
|||||
290 TG.GCAGACGAACAGAACGTTGAGCGCTGCACAGCACCCAGGCTGCGGG 338
|||||
56 ..... 56
|||||
339 GGTACCGGAGCCCYACCCCTGGCCCTGAGGGAGAGGGAAGTCAGGGCCGG 388
|||||
57 ..... ValasnGlyPheAsnThrLeu 63
|||||
389 GGCTGCGCCCAATCCTGTCTCTGTGCGCAGTGAACGGATTCAACACCCTT 438
|||||
64 GlyGluasnIlealaaspasnGlyGlyValArgGlnAlaTyrLysAlaTyr 80
|||||
439 GGGGAAACAATTCTGTACAAACGGAGGGGTGCGGCAGGCCTATAAGGCCTA 488
|||||
80 rLeuLysTrrpMetAlaGluGlyGlyLysaspGlnGlnLeuProGlyLeuA 97
|||||
489 CCTCAAGTGGATGGCAGAGGGTGGCAAGGACCAAGCAGCTGCCCGGCCCTGG 538
|||||
97 sPLeuthrHisGluGlnLeuPhePheIleAsnTyrAlaGlnValTrrpCys 113
|||||
539 ATCTCACCATGAGCAGCTCTTCTTCATCACTATGCCCCCAGGTGTGGTGC 588
|||||
114 Gly 114
|||||
589 GGG 591

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/va2000.DAT:AAA63
seq_documentation_block:
ID AAA63763 standard; cDNA; 2925 BP.

```

seq_name: /SIDSl/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:AAA63763

seq_documentation_block:	
ID	AAA63763 standard; cDNA; 2925 BP.
XX	
XX	AAA63763;
XX	
DT	04-DEC-2000 (first entry)
DT	
XX	
DE	cDNA encoding neutral endopeptidase metalloproteinase-like enzyme NL-1.
XX	
KW	Neprilysin; neutral endopeptidase metalloproteinase-like enzyme;
KW	NEP-like enzyme; protein production; protein secretion;
KW	neurological disease; Alzheimer's disease; pain; psychiatric disorder;
KW	fertility; bone disease; abnormal phosphate metabolism; ss.
XX	
OS	Mus sp.
XX	
Key	Location/Qualifiers
FFH	332..2629
CDS	/*tag= a
FT	/product= "neutral endopeptidase metalloproteinase-like
FT	enzyme NL-1"
FT	
XX	
XX	WO2000047750-A2.
XX	
XX	17-AUG-2000.
XX	
XX	11-FEB-2000; 2000WO-CA00147.
PPF	
XX	
PR	11-FEB-1999; 99CA-2260376.
XX	
PA	(UYMO-) UNIV MONTREAL.
XX	
PI	Desgroseillers L, Boileau G;
XX	
XX	WPI; 2000-549148/50.
DR	P-PSDB; AAB08130.
DR	

xx Rat; membrane metalloprotease; neprilysine II; NEPII; inactivation;
KW neuron; hormone; peptide messenger; inhibitor; detection; disorder;
KW cardiovascular disease; neurodegenerative disease; growth disorder;
KW hypothalamic-hypophyseal axis; endocrine disorder; ds.
XX
OS Rattus rattus.
XX
PN FR2777291-A1.
XX
PD 15-OCT-1999.
XX
PF 08-APR-1998; 98FR-0004389.
XX
PR 08-APR-1998; 98FR-0004389.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Ouimet T, Gros C, Haret C, Bonhomme MC, Facchinetti P;
PI Schwartz JC;
XX
XR WPI: 1999-593429/51.
DR P-PSDB; AAY44177.
XX
XX New membrane metalloprotease NEP II, involved in proteolysis of
PT neuronal and hormonal peptides, used to screen for inhibitors.
PT potentially useful for treating e.g. cardiovascular disease .
XX
XX Claim 2; Page 12-16; 29pp; French.
XX
XX This sequence represents the gene for the rat membrane metalloprotease
CC designated neprilysine II (NEPII), which is involved in (in)activation
CC of neuronal and hormonal peptide messengers. NEPII is used to screen
CC for specific substrates (used to detect NEPII in cells and tissues) or
CC inhibitors, which can also be used to detect NEPII or for treatment of
CC disorders related to peptidergic signalling in which NEPII is involved,
CC e.g. cardiovascular or neurodegenerative diseases; growth disorders of
CC endocrine origin; disturbances of the hypothalamic-hypophyseal axis or
CC endocrine disorders.
XX
SQ Sequence 2765 BP; 684 A; 735 C; 787 G; 559 T; 0 other;

alignment_scores:
Quality: 578.00 Length: 114
Ratio: 5.115 Gaps: 0
Percent Similarity: 99.123 Percent Identity: 85.088
alignment_block:
US-09-647-780A-4 x AA228810
Align seg 1/1 to: AA228810 from: 1 to: 2765
1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPhesAspLy 17
1925 GGACACGAGATCACACGGCTTTGATGATACGCTCGGAACCTTTGACAA 1974
17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPhesSerThrGlnHisPheA 34
1975 GAATGGCAACATGCTGGACTGGTGAGCAACTTCTCGGCCGCGCACTTCC 2024
34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
2025 GACAGCAGTCACAGTGTATGATTATCATAGTACAGCAACTTCTCTGGGAA 2074
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnII 67
2075 CTACGACACACACAGATGTGACGGATTTCAGCACCCTCGGGGAGAACAT 2124
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
2125 CGCGCAACACGGCGTGTGCGGAGGCATACAAGCTTACCTACAGTGGC 2174

84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
2175 TAGCTGAGGCGCGAGACACAGAGACTGCCGGGACTGAACCTGACCTAT 2224
101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114
2225 GCTCAGCTTTTCTTCATTAACATATGCCAGGTGTGTGTGG 2266
seq_name: /SIDS1/gcdata/geneseq/geneseq-nbml/NA1988.DAT.AAN81690
seq_documentation_block:
ID AAN81690 standard; DNA; 3369 BP.
XX
AC AAN81690;
XX
DT 23-NOV-1990 (first entry)
XX
DE Atrial natriuretic polypeptide binding polypeptide encoding gene.
XX
KW Diuretic; atrium cardis; atrial natriuretic peptide binding;
KW hypotensive action; ss.
XX
FH Key Location/Qualifiers
CDS 23..2272
FT /*tag= a
FT /product=atrial natriuretic polypeptide binding peptide
XX
PN JP63079598-A.
XX
PD 09-APR-1988.
XX
PF 22-SEP-1986; 86JP-0222192.
XX
PR 22-SEP-1986; 86JP-0222192.
XX
PA (SUNR) SUNTORY LTD.
XX
XX WPI: 1988-137132/20.
DR P-PSDB; AAP82940.
XX
XX Novel polypeptide, with diuretic action - is obtd. from atrium cardis
PT of mammals showing specific bond to atrial natriuretic polypeptide
PT and gene coding it.
XX
PS Disclosure; ; p; Japanese.
XX
XX The peptide encoded by the gene specifically binds to atrial
CC natriuretic polypeptide (ANP) and has diuretic (partic. natriuretic)
CC and hypotensive action.
CC See also AAN81690-93 and AAP81282-P81309.
XX
SQ Sequence 3369 BP; 1110 A; 595 C; 691 G; 973 T; 0 other;

alignment_scores:
Quality: 471.00 Length: 114
Ratio: 4.402 Gaps: 0
Percent Similarity: 93.860 Percent Identity: 66.667
alignment_block:
US-09-647-780A-4 x AAN81690
Align seg 1/1 to: AAN81690 from: 1 to: 3369
1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPhesAspLy 17
1769 GGACACGAAATCACCCATGGCTTCGATGACAAATGCAGAAATTTTAACAA 1818
17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPhesSerThrGlnHisPheA 34
1819 GGATGGAGACCTTGTGACTGTGTGACTGACCAACAGCTGCAAAATAATTTTA 1868
34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50

```

1869 AAGAACAATCCCAATGATGGTGTACCACTATGGAACACTCTCTCTGGGAC 1918
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnI1 67
1919 CTAGCAGGTGACAGCATCTCAATGGAATTAATACACTGGGAGAAACAT 1968
67 eAlaAspAsnGlyGlyValAlaGlnAlaTyrLysAlaTyrLeuLysTrpM 84
1969 TGCTGATACCGGTGTATGGCCAGCATACAGACCTTCAAAATATG 2018
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
2019 TTAAAGAAGATGGTGAAGAAAAATTAATCTCTGGGACTTGACCTAAATCAT 2068
101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114
2069 AAACAACATGTTCTTCTTGAACCTTCGCCAGGTGTGTGTGGG 2110

```

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.AAN80952

seq_documentation_block:

ID AAN80952 standard; DNA; 3181 BP.

XX AAN80952;

AC AC

DT 15-NOV-1990 (first entry)

XX Enkephalinase gene (human).

XX Enkephalinase; immunosuppressant; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 3..2249

FT /tag= a

FT /product=enkephalinase

XX EP272928-A.

XX 23-DEC-1987; 87EP-0311365;

XX 12-JAN-1987; 86US-0946566.

XX 29-JUN-1988; 87US-0002478.

XX (GETH) GENENTECH INC.

XX Malfroy-Camine B, Scofield PR;

XX WPI; 1988-177004/26.

XX Deoxyribonucleic acid isolates coding for enkephalinase - and prodn.

XX of enkephalinase, used in therapeutic applications in diagnostic

XX immunoassay and as immunosuppressants.

XX Disclosure; ; p; English.

XX The DNA was sequenced from clones isolated from a lambda gt10

XX library constructed from poly(A+) human placental RNA. At posn.

XX 1413 one clone, lambdaH7 was observed to have a G (coding for Ala);

XX in clone lambdaH5 this was found to be an A (Thr). Since the former

XX is identical to the rat amino acid at posn. 465, the latter

XX probably represents an error of the reverse transcriptase synthesis

XX of the mRNA. The DNA can be ligated into an expression vector for

XX prodn. of the enkephalinase protein in a host. The recombinant

XX protein is useful in the treatment of various pathological disor

XX ders associated with endogenous peptides such as tachykinins and

XX kinins, eg. acute inflammation, and hyperimmune responses. It's

XX also useful in diagnostic immunoassays, and as immunosuppressants

XX by virtue of their ability to digest chemotactic molecules.

CC This sequence is also published in EP-272929.

CC See also AAN80950, AAN80951 and AAN82261.

XX Sequence 3181 BP; 1056 A; 584 C; 653 G; 887 T; 1 other;

alignment_scores:

Quality: 469.00 Length: 114
Ratio: 4.383 Gaps: 0
Percent Similarity: 93.860 Percent Identity: 66.667

alignment_block:

US-09-647-780A-4 x AAN80952

Align seg 1/1 to: AAN80952 from: 1 to: 3181

```

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAspLy 17
|||||
1743 GGACACGAAATCACCCTCGCTCGATGACAATGGCAGAACTTTAACAA 1792
17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
|||||
1793 AGATGGAGACCTCGTTGACTGTGGTCAACAGCTGCAAGTAACCTTA 1842
34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
|||||
1843 AGAGCAATCCAGTCATGGTGTATCAGTATGGAATTAATACACTGGGAAACAT 1942
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnI1 67
|||||
1893 CTGGCAGGTGGACAGCACCTTAATGGAATTAATACACTGGGAAACAT 1942
67 eAlaAspAsnGlyGlyValAlaGlnAlaTyrLysAlaTyrLeuLysTrpM 84
|||||
1943 TGCTGATAATGGAGTCTTGGTCAAGCATACAGAGCCTATCAGAATTATA 1992
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
|||||
1993 TTAAAGAAGATGGCGAAGAAAAATTAATCTCTGGGACTTGACCTAAATCAC 2042
101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114
|||||
2043 AAACAACATATTTCTTGAACCTTGCACAGGTGTGTGTGGA 2084

```

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.AAQ62678

seq_documentation_block:

ID AAQ62678 standard; cDNA; 3181 BP.

XX AAQ62678;

XX 05-JAN-1995 (first entry)

XX Human enkephalinase cDNA.

XX Human; enkephalinase; neutral endopeptidase; kidney; intestine; kinin;

XX kidney brush border neutral proteinase; pituitary; brain; lymph nodes;

XX neutrophils; enkephalins; tachykinins; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 3..2249 /tag= a

FT mat_peptide 21..2246 /tag= b

FT allele 1413 /tag= c

FT /note= "Nucleotide is A in a different clone

FT changing the codon to Thr"

XX EP596355-A.

XX


```

264 CGCGACACACGGGGGCGCTCAAGCGGGCGCTATCGGGCTACCAAGACTGGG 313
      84 etAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      314 TCAAGAGATGGGGCTGACACACACTGCCACCTCGGTCTCACCACAC 363
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      101 GluGlnLeuPhePheIleAsnTrpAlaGlnValTrpCysGlyCysLys 116
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      364 AACCAAGCTCTCTCTCTGAGTTTGGACAGGCTGGGTGTTCCGTCGCGC 411
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-646-273-22

```

```

seq documentation_block:
; Sequence 22, Application US/08646273
; Patent No. 6066502
;
; GENERAL INFORMATION:
; APPLICANT: Kroeger, Burkhard, Seulberger, Harald, Meyer, Thomas, Schmidt,
; APPLICANT: Martin, Jacob, Elard, Otter, Rainer, Subkowski, Thomas, Hillen, Heinz
; TITLE OF INVENTION: Endothelin Converting Enzyme (ECE).
;
; NUMBER OF SEQUENCES: 36
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Keil & Weinkauf
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
; COMPUTER: IBM AT-compatible, 80486 processor
; OPERATING SYSTEM: MS-DOS version 6.0
; SOFTWARE: Wordperfect version 5.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,273
; FILING DATE: 16-NOV-1994
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/03706
; FILING DATE: 11-NOV-1994
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2129 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA for mRNA
;
; US-08-646-273-22

```

```
alignment_scores:      Quality: 374.50      Length: 116  
                       Ratio: 3.901       Gaps: 1  
Percent Similarity: 82.759 Percent Identity: 54.310  
  
alignment_block:  
US-09-647-780A-4 x US-08-646-273-22  
  
Align seg 1/1 to: US-08-646-273-22 from: 1 to: 2129  
  
1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAspLy 17  
|||||.....  
1632 GCCACGAGCTCAGTCATCTTTGATGATCAAGGCCGAGAGTACCACAA 1681  
  
17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34  
|||||.....  
1692 GGATGGGAACCTCCCGCCCTGTGGGAAGAACTCGTCGGTGAGGCCTTCA 1731  
  
34 rGluGlnSerGluCySMetIleTyrglnTyrglyAsnTyrserTrpasp 50  
.....  
1732 ACAGACAGCCGGTCATGGTGGACCATACGGCAACTATACC..... 1775  
  
51 LeuAlaaspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluasnII 67  
.....
```

```

1776 ...GTGAACGGGGAGCCGGTGAACACGGCCGGCACCCCTCGCGGAAACAT 1822
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
1823 CGCCGACAAACGGGGCCCTCAAGCGCGCCTATCGGGCTACCAAGATGGG 1872
84 eAlaGluGlyCysLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
1873 TCAAGAAGAATGGCGCTGAGCAGACACTGCCACCCCTGGTGTCAACAAC 1922
101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGlyCysLys 116
1923 AACCAAGCTTTCTTCTCTGAGTTTGGTTGGACAGGTCCTGGTGTTCCTCGCG 1970
seq name: /can2_6/otodata/1/ina/6A_COMB.seq;seq:US-08-646-273-29

```

```

seq_documentation_block:
; Sequence 29, Application US/08646273
; Patent No. 6066502
; GENERAL INFORMATION:
; APPLICANT: Kroeger, Burkhard, Seulberger, Harald, Meyer, TH
; APPLICANT: Martin, Jacob, Elard, Otter, Rainer, Subkowski,
; TITLE OF INVENTION: Endothelin Converting Enzyme (ECE).
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Keil & Weinkauff
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 MB storage.
; COMPUTER: IBM AT-compatible, 80486 processor
; OPERATING SYSTEM: MS-DOS version 6.0
; SOFTWARE: Wordperfect version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,273
; FILING DATE: 16-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/03706
; FILING DATE: 11-NOV-1994
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2314 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA for mRNA
; US-08-646-273-29

```

```

alignment_scores:
  Quality: 374.50      Length: 116
  Ratio: 3.901        Gaps: 1
  Percent Similarity: 82.759  Percent Identity: 54.310

alignment_block:
  US-09-647-780A-4 x US-08-646-273-29  ..

Align seg 1/1 to: US-08-646-273-29  from: 1 to: 2314

1  GLYHISGLUILETHRHSGLYPHEASPPASPNGLYARGASNPHEASPLY 17
      |||.....|.....|.....|.....|.....|.....|
1806  GGCACGACCTGACTCATCTGCTTTGATGATCAAGCGCGAGTACGACAA 1855

17  sASnGLYASnMetMetAspTrpSerASnPSheSerThrGlnHISphea 34
      |||.....|.....|.....|.....|.....|.....|
1856  GGATGGGAACCTCCGGCCCTCGTGTGGAAGAAGACTCGTCCTGGAGCGCTCA 1905

34  rgGluGlnSerGluCysMetIleTyrGlnTyrGlnTyrSerTrpAsp 50
      |||.....|.....|.....|.....|.....|.....|

```

1906 AGCAGCAGACCGCGTCATGGTGGAGCAGTACGGCAACTATATAGC..... 1949
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnI1 67
1950GTGAACGGGAGCGCGGTGAACCGCGGACACCCCTCGGCGAAACAT 1996
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
1997 GCGCGACAACGGGCGCTCAAGGCGGCGCTATCGGCGCTACCAAGAACTGGG 2046
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
2047 TCAAGAAGATGGGCTGAGCAGACACTGCCACCCCTGGGTCTCCACCAAC 2096
101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGlyCysLys 116
2097 AACCAAGCTCTTCTCTGAGTTTCCACAGGCTGTGTTCGGTCCGCGC 2144
seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-289-112-1

seq_documentation_block:
; Sequence 1, Application US/08289112
; Patent No. 568640
; GENERAL INFORMATION:
; APPLICANT: Yanagisawa, Masashi
; TITLE OF INVENTION: Endothelin Converting Enzyme-1: A
; TITLE OF INVENTION: Membrane-Bound Metalloprotease That Catalyzes The
; TITLE OF INVENTION: Proteolytic Activation of Big Endothelin-1
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/289,112
; FILING DATE: 10-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSD:414/PAR
; TELEPHONE: 713-787-1400
; TELEFAX: 713-789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2889 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 118..2391
; US-08-289-112-1

alignment_scores:
Quality: 370.50 Length: 116
Ratio: 3.859
Percent Similarity: 82.759 Percent Identity: 53.448

alignment_block:
US-09-647-780A-4 x US-08-289-112-1

Align seg 1/1 to: US-08-289-112-1 from: 1 to: 2889
1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAspLys 17
1897 GCCCAGGAGCTGACATCATCTTTGATGATCAAGGCCGAGAGTACGACAA 1946
17 sAsnGlyAsnMetMetAspTyrTrpSerAsnPheSerThrGlnHisPheA 34
1947 GGATGGGAACCTCCGGCCCTGGTGGGAAGAACTCGTCGCGGAGCGGTTC 1996
34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
1997 AGCAGCAGACCGCGTGCATGGTGGAGCAGTACGGCAACTATAGC..... 2040
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnI1 67
2041GTGAACGGGAGCGCGGTGAACCGCGGACACCCCTCGGCGAAACAT 2087
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
2088 CCGCGACAACGGGCGCTCAAGGCGGCGCTATCGGCGCTACCAAGAACTGGG 2137
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
2138 TCAAGAAGATGGGCTGAGCAGACACTGCCACCCCTGGGTCTCCACCAAC 2187
101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGlyCysLys 116
2188 AACCAAGCTCTTCTCTGAGTTTCCACAGGCTGTGTTCGGTCCGCGC 2235
seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-646-273-24

seq_documentation_block:
; Sequence 24, Application US/08646273
; Patent No. 6066502
; GENERAL INFORMATION:
; APPLICANT: Kroeger, Burkhard, Seulberger, Harald, Meyer, Thomas, Schmidt,
; APPLICANT: Martin, Jacob, Elard, Otter, Rainer, Subkowski, Thomas, Hillen, Hein
; TITLE OF INVENTION: Endothelin Converting Enzyme (ECE)
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Keil & Weinkauff
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
; COMPUTER: IBM AT-compatible, 80486 processor
; OPERATING SYSTEM: MS-DOS version 6.0
; SOFTWARE: Wordperfect version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,273
; FILING DATE: 16-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/03706
; FILING DATE: 11-NOV-1994
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2533 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA for mRNA
; US-08-646-273-24

alignment_scores:
Quality: 367.50 Length: 116
Ratio: 3.750
Gaps: 1

```

alignment_scores:
  Quality: 367.50      Length: 116
  Ratio: 3.750        Gaps: 1
  Percent Similarity: 84.483  Percent Identity: 53.448

alignment_block:
  US-09-647-780A-4 x US-08-646-273-35

Align seg 1/1 to: US-08-646-273-35 from: 1 to: 2720

1 GlyHisGIuIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17
1802 GGCCATGAGCTGACATCACTGCTTTGATCATCAAGGACGGGAGTACGACA 1851

17 sAsnGlyAsnMetMetAspTrrpTrrpSerAsnPheSerThrGlnHisPheA 34
1852 GGACGGGGAACCTCCGGCCATGCTGTGAAGAAGACTCATCCGTGGAGGCCTTCA 1901

34 rgGlulnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
1902 AGCGTCACACCGAGTGCATGCTGTAGACAGTAGTACAGAACTACAGC..... 1945

51 LeuAlaAspGluInAsnValAsnGlyPheAsnThrLeuGlyClnAsnIrl 67
1946 ...GTGAACGGGGAGCCCGGTGAACGGGGGACACCCCTGGGGAGAACAT 1992

67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
1993 GCGCCACAAACGGGGGTCTCAAGCGCGCCTATCGGCTTACCAGCAACTGGG 2042

84 etAlaGluClyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
2043 TGAAGAAGAAACGGGGCTGAGCACTCGCTCCCAACCTGGGCGCTCACCAAT 2092

101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGlyCysLys 116
2093 AACCAGCTCTTCCTCGTGGCTTTGCACAGGTCTGGTGTCTCGCTCGCT 2140

seq name: /cgn2 6/ptodata/1/ina/5A_COMB.seq:US-08-574-763-1

```

```

seq_documentation_block:
; Sequence 1, Application us/08574763.
; Patent No. 5736376
; GENERAL INFORMATION:
; APPLICANT: Yanagisawa, Masashi
; TITLE OF INVENTION: RECOMBINANT ENDOTHELIN CONVERTING
; TITLE OF INVENTION: ENZYME-2 AND ITS USE IN ECE INHIBITOR SCREENING
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,763
; FILING DATE: Concurrently herewith
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Corder, Timothy S.
; REGISTRATION NUMBER: 38,414
; REFERENCE/DOCKET NUMBER: UTXD:472
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-5777
; INFORMATION FOR SEQ ID NO: 1:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 3291 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 116..2476
US-08-574-763-1

alignment_scores:
Quality: 339.50 Length: 116
Ratio: 3.574 Gaps: 1
Percent Similarity: 81.897 Percent Identity: 50.000

alignment_block:
US-09-647-780A-4 x US-08-574-763-1

Align seg 1/1 to: US-08-574-763-1 from: 1 to: 3291

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAspLy 17
|||||
1982 GCCACGAGTTGACACATGCTTTCATCACCAGGCGCGAGTATGACAA 2031
17 sasnGlyAsnMetMetAspTTrpSerAsnPheSerThrGlnHisPheA 34
|||||
2032 GGAAGGAAATCGCGCGCTGTGGCAGATGAGTCGTGGCAGCCTCC 2081
34 rgluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
|||||
2082 GGAACACACGCGCTGATAGAGGAGCAGTACAGCCAGTACCAG 2125
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnI 67
|||||
2126 ...GTCAACGGGGAGAGCACTCAACGGCGGCCAGACACTGGGGGAGAACAT 2172
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
|||||
2173 TGCCGCAATGGGGGCTTAGGCTGCTTACACCTTACAAAGCATGGC 2222
2223 TAAGAAAGCATGGGAGGAGCAGCAGCTGCCAGCTGTGGACTCACCAAC 2272
101 GluGlnLeuPheIleAsnTyrAlaGlnValTrpCysGlyCysLys 116
|||||
2273 CACCAGCTCTCTTTGTGGGATTTGCCAGGTGTGGTGTCTGCTGCTCCG 2320

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-08-646-273-18

seq_documentation_block:

Sequence 18, Application US/08646273
Patent No. 6065502
GENERAL INFORMATION:
APPLICANT: Kroeger, Burkhard, Seulberger, Harald, Meyer, Thomas, Schmidt,
APPLICANT: Martin, Jacob, Elard, Otter, Rainer, Subkowski, Thomas, Hillen, Heinz.
TITLE OF INVENTION: Endothelin Converting Enzyme (ECE).
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Keil & Weinkauf
STREET: 1101 Connecticut Avenue
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
COMPUTER: IBM AT-compatible, 80486 processor
OPERATING SYSTEM: MS-DOS version 6.0
SOFTWARE: WordPerfect version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646, 273

FILING DATE: 16-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/03706
FILING DATE: 11-NOV-1994
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1703 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA for mRNA
US-08-646-273-18

alignment_scores:
Quality: 336.50 Length: 106
Ratio: 3.868 Gaps: 1
Percent Similarity: 82.075 Percent Identity: 54.717

alignment_block:

US-09-647-780A-4 x US-08-646-273-18

Align seg 1/1 to: US-08-646-273-18 from: 1 to: 1703

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAspLy 17
|||||
1394 GCCCAGGAGCTGACTCATCTTTTGTATGATCAAGCGCGAGTACGACAA 1443
17 sasnGlyAsnMetMetAspTTrpSerAsnPheSerThrGlnHisPheA 34
|||||
1444 GATGGGAACCTCCGGCCCTGTGGGAGAACTCGTCCGTGGAGCGCTTCA 1493
34 rgluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
|||||
1494 AGCAGCAGACCGCGTGCATGCTGGTGGAGCAGTACGGCAACTATAGC 1537
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnI 67
|||||
1538 ...GTGAACGGGGAGCGCGTGAACGGCGGCACACCTCGCGGAAACAT 1584
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
|||||
1585 CCCCACAAACGGGGCGCTCAGCGCGCTATCGGCGCTACCGAACTGGG 1634
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
|||||
1635 TCAAGAGAATGGGCTGAGCAGACACTGCCACCTGGGTCTCACCACAC 1684
101 GluGlnLeuPhePheIle 106
|||||

1685 AACCAGCTCTCTCTCTG 1702

seq_name: /cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-305-640-1

seq_documentation_block:

Sequence 1, Application US/09305640B
Patent No. 6255468
GENERAL INFORMATION:
APPLICANT: SmithKline Beecham plc
TITLE OF INVENTION: No. 6255468el Compounds
FILE REFERENCE: GP30124
CURRENT APPLICATION NUMBER: US/09/305,640B
CURRENT FILING DATE: 1999-05-05
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2560
TYPE: DNA
ORGANISM: Homo sapiens
US-09-305-640-1

```
alignment_scores:
  Quality: 318.50      Length: 113
  Ratio: 3.703        Gaps: 1
  Percent Similarity: 76.106  Percent Identity: 49.558

alignment_block:
US-09-647-780A-4 x US-09-305-640-1
Align seg 1/1 to: US-09-305-640-1 from: 1 to: 2560

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2056 GGACATGAGCTGACCCACGCGTACGACGCTGGGGGCCGATGACCG 2105

17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2106 CTCAGGGAACCTGCTGCACTGGTGACGAGCGCCCTACAGCCGCTTCC 2155

34 rGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50
|||||:|||||:|||||:|||||:|||||:|||||:
2156 TCGGAAGCGTGAAGTGCATGCTGCTCTATGACAACTTCACT ..... 2199

51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsn 67
|||||:|||||:|||||:|||||:|||||:|||||:
2200 ...GCTACACACCGCGGTGAACGGAACACACGCTGGGAGAACAT 2246

67 eAlaAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
|||||:|||||:|||||:|||||:|||||:|||||:
2247 GCGCATATGGCGCCCTCAAGCTGGCCTTACCACGCTATCAGAAAGTGG 2296

84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
|||||:|||||:|||||:|||||:|||||:|||||:
2297 TCGGGAGCAGCGCCGACGAGACCCACTTCCCGGCTCAAGTACACACAT 2346

101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCys 113
|||||:|||||:|||||:|||||:|||||:|||||:
2347 GACCAGCTCTTCTTCACTTGCCTTCCCGCAGAACTGGTGC 2385

seq_name: /cgn2_6/ptodata/1/ina/6B_COMB.seq:US-09-305-640-3

seq_documentation_block:
; Sequence 3, Application US/09305640B
; Patent No. 6255468
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham plc
; TITLE OF INVENTION: No. 6255468el Compounds
; FILE REFERENCE: GP30124
; CURRENT APPLICATION NUMBER: US/09/305,640B
; CURRENT FILING DATE: 1999-05-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version. 3.0
; SEQ ID NO 3
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (404)(437)(460)(490)(520)(555)(561)(595)
US-09-305-640-3

alignment_scores:
  Quality: 112.50      Length: 83
  Ratio: 1.907        Gaps: 5
  Percent Similarity: 71.084  Percent Identity: 42.169

alignment_block:
US-09-647-780A-4 x US-09-305-640-3
Align seg 1/1 to: US-09-305-640-3 from: 1 to: 606

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
305 GGACATGAGCTGACCCACGCGTACGACGAGCTGGGGGCCGAGTATGACC 354
17 ysAsnGlyAsnMetMetAspTrpTrp...SerAsnPheSerThrGlnHis 32
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
355 GCTCAGGGAACCTGCTGCACTGGTGAGCGGAGGAGCTTCTTACAGCCGN 404
32 spHeArgGluGlnSerGluCysMetIle...TyrGlnTyrGlyAsnTyrSe 48
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
405 TTCTCTGCGAAGGCTGAGTGCATGCTCCCTNTTTTATGGACAATTC 454
48 rTrpAspLeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyG 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
455 A.....ATGNTTACAAACGCGGTGAACGGGAA.ACACANGTTTGGG 497
65 luAsnIleAlaAspAsnGlyGlyValArgGlnAlaTyrLysAla 79
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
498 AGAACATCGCAGTATGGGGCGGNCCTTAAGTTGGCTTACCACGCT 541

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-258-188-1

seq_documentation_block:
; Sequence 1, Application US/08258188
; Patent No. 5475098
; GENERAL INFORMATION:
; APPLICANT: HALL, Robert H.
; APPLICANT: XU, Jian Guo
; TITLE OF INVENTION: A NEW AND DISTINCTIVE DNA SEQUENCE OF E.
; TITLE OF INVENTION: coli O157:H7 AND ITS USE FOR THE RAPID, SENSITIVE AND
; TITLE OF INVENTION: SPECIFIC DETECTION OF O157:H7 AND OTHER ENTEROHEMORRHAGIC
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/258,188
; FILING DATE: 14-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 15280206, DHHSE135940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2278 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (plasmid)
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..2275
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2278
; OTHER INFORMATION: /product= "Peptide encoded by the
; OTHER INFORMATION: hlyA gene"
; OTHER INFORMATION: /standard_name= "Nucleic acid sequence of hlyA
; OTHER INFORMATION: gene"
US-08-258-188-1
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2278 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (plasmid)
FEATURE:
NAME/KEY: CDS
LOCATION: 2..2275
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..2278
OTHER INFORMATION: /product= "peptide encoded by the
OTHER INFORMATION: hlyA gene"
OTHER INFORMATION: /standard_name= "Nucleic acid sequence
OTHER INFORMATION: of hlyA gene"
US-08-526-813-1

alignment_scores:
Quality: 73.00 Length: 91.
Ratio: 1.521 Gaps: 4
Percent Similarity: 52.747 Percent Identity: 24.176

alignment_block:
US-09-647-780A-4 x US-08-526-813-1
Align seg 1/1 to: US-08-526-813-1 from: 1 to: 2278

seq_name: /cgn2_6/ptodata/1/ina/3A_COMB.seq:US-08-526-813-1

seq_documentation_block:
Sequence 1, Application US/08526813
Patent No. 5756293
GENERAL INFORMATION:
APPLICANT: Hall, Robert H.
APPLICANT: Xu, Jian Guo
TITLE OF INVENTION: A New and Distinctive DNA Sequence of E.
Patent No. 5756293
TITLE OF INVENTION: coli O157:H7 and its Use for the Rapid, Sensitive and
TITLE OF INVENTION: Specific Detection of O157:H7 and Other Enterohemorrhagic
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/526,813
FILING DATE: 11-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,188
FILING DATE: 14-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DOCKET NUMBER: 15280-206-1

10 AspAsnGlyArgAsnPheAspLysAsnGly..... 19
542 GAGCATGCGCAAAATATTTTGAGAAATGGATATGACGCAAGACATGCTGC 591
20AsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisP 33
592 GTTTTGTAGAGACACTCTGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTG 641
33 heArgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrp 49
642 CAGTAGAAGAGCAGCTGCGCAATAACCCAGCAACAT.....TGG 679
50 AspLeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGlu.. 65
680 GAT.....GAGAAGATCGGTGAACAT..... 772
66AsnIleAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrL 81
700 TGCAGGCGATAACCCGTAATGCTGATGCTGATGCTGATGCTGATGCTGATG 749
81 euLysTrpMetAlaGluGlyGly 88
750 TTAATTATCTGGAATGGAGG 772

seq_name: /cgn2_6/ptodata/1/ina/3A_COMB.seq:US-08-526-813-1

seq_documentation_block:
Sequence 1, Application US/08526813
Patent No. 5756293
GENERAL INFORMATION:
APPLICANT: Hall, Robert H.
APPLICANT: Xu, Jian Guo
TITLE OF INVENTION: A New and Distinctive DNA Sequence of E.
Patent No. 5756293
TITLE OF INVENTION: coli O157:H7 and its Use for the Rapid, Sensitive and
TITLE OF INVENTION: Specific Detection of O157:H7 and Other Enterohemorrhagic
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/526,813
FILING DATE: 11-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,188
FILING DATE: 14-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DOCKET NUMBER: 15280-206-1

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2278 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (plasmid)
FEATURE:
NAME/KEY: CDS
LOCATION: 2..2275
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..2278
OTHER INFORMATION: /product= "peptide encoded by the
OTHER INFORMATION: hlyA gene"
OTHER INFORMATION: /standard_name= "Nucleic acid sequence
OTHER INFORMATION: of hlyA gene"
US-08-526-813-1

alignment_scores:
Quality: 73.00 Length: 91.
Ratio: 1.521 Gaps: 4
Percent Similarity: 52.747 Percent Identity: 24.176

alignment_block:
US-09-647-780A-4 x US-08-526-813-1
Align seg 1/1 to: US-08-526-813-1 from: 1 to: 2278

seq_name: /cgn2_6/ptodata/1/ina/3A_COMB.seq:US-08-526-813-1

seq_documentation_block:
Sequence 1, Application US/08526813
Patent No. 5756293
GENERAL INFORMATION:
APPLICANT: Hall, Robert H.
APPLICANT: Xu, Jian Guo
TITLE OF INVENTION: A New and Distinctive DNA Sequence of E.
Patent No. 5756293
TITLE OF INVENTION: coli O157:H7 and its Use for the Rapid, Sensitive and
TITLE OF INVENTION: Specific Detection of O157:H7 and Other Enterohemorrhagic
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/526,813
FILING DATE: 11-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,188
FILING DATE: 14-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DOCKET NUMBER: 15280-206-1

seq_name: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:PCT-US95-08554-1

seq_documentation_block:
Sequence 1, Application PC/TUS9508554
GENERAL INFORMATION:
APPLICANT: HALL, Robert H.
APPLICANT: XU, Jian Guo
TITLE OF INVENTION: A NEW AND DISTINCTIVE DNA SEQUENCE OF E.
TITLE OF INVENTION: coli O157:H7 AND ITS USE FOR THE RAPID,
TITLE OF INVENTION: SPECIFIC DETECTION OF O157:H7 AND OTHER
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco


```
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08554
FILING DATE: 14-JUN-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DOCKET NUMBER: 15280206.DHHSE135940
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2278 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (plasmid)
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 2..2275
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1..2278
OTHER INFORMATION: /product= "Peptide encoded by the
OTHER INFORMATION: hlyA gene"
OTHER INFORMATION: /standard_name= "Nucleic acid sequence of hlyA"
PCT-US95-08554-1

alignment_scores:
Quality: 73.00 Length: 91
Ratio: 1.521 Gaps: 4
Percent Similarity: 52.747 Percent Identity: 24.176

alignment_block:
US-09-647-780A-4 x PCT-US95-08554-1
Align seg 1/1 to: PCT-US95-08554-1 from: 1 to: 2278
10 ASPASNGLYARGASNPHASPLYSASNGLY..... 19
:|||||:|||||:|||||:|||||:
542 GAGCATGCGCAAAATATTATTTGAGATGGATATGACGCAAGACATGCTGC 591
20 .....AsnMetMetAspTTPTrpSerAsnPheserThrGlnHisP 33
592 GTTTTATAGAGACTCTCTGCTTTTCTGCTTGATTTTCTCGTCAGCATG 641
642 CAGTAGAAGAGCAGTCGCAATAACCCAGCAACAT.....TGG 679
50 AspLeuAlaaspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGlu.. 65
|||:|||||:|||||:|||||:
680 GAT.....GAGAAGATCGGTGAAC 699
66 .....AsnIleAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrL 81
:|||||:|||||:|||||:|||||:
700 TGCAGGCATACCCGTAATGCTGATCGCAGTCAGTGGTAAGGCATATA 749
81 euLysTrpMetAlaGluGlyGly 88
:|||||:|||||:|||||:
750 TTAATTATCTGGAATAATGGAGG 772

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-337-268A-54
seq_documentation_block:
Sequence 54, Application US/08337268A
Patent No. 5589336
GENERAL INFORMATION:
APPLICANT: Lee, Soohae
APPLICANT: Redman, Colvin L.
TITLE OF INVENTION: Diagnostic Method and Kit for
TITLE OF INVENTION: Determining Kell Blood Group
TITLE OF INVENTION: Genotype
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ronald J. Baron, Esq.
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: USA
ZIP: 11753
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/337,268A
FILING DATE: 11-OCT-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA: No. 5589336e
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 454-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
HYPOTHETICAL: no
US-08-337-268A-54

alignment_scores:
Quality: 72.00 Length: 37
Ratio: 2.880 Gaps: 0
Percent Similarity: 67.568 Percent Identity: 43.243

alignment_block:
US-09-647-780A-4 x US-08-337-268A-54
Align seg 1/1 to: US-08-337-268A-54 from: 1 to: 146
78 LysAlaTyrLeuLysTrpMetAlaGluGlyGlyLysAspGlnGlnLeuPr 94
:|||||:|||||:|||||:|||||:
23 CAGCATACACACAGAGCTGTTAGCGCACCATGGGAGACTGCTCTGCC 72
94 oGlyLeuAspLeuThrHisGluGlnLeuPhePheIleAsnTyrAlaGlnV 111
:|||||:|||||:|||||:|||||:
73 CAGCTGGACCTCAGCCCCCAGCAGATCTTTTCGAAGCTATGCCAGG 122
111 alTrpCysGly 114
|||:|||||:
123 TAGGCAGCGGC 133

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-484-570A-54
```

```
seq_documentation_block:
; Sequence 54, Application US/08484570A
; Patent No. 5804379
; GENERAL INFORMATION:
; APPLICANT: Lee, Soohae
; APPLICANT: Redman, Colvin L.
; TITLE OF INVENTION: Diagnostic Method and Kit for
; TITLE OF INVENTION: Determining Kell Blood Group
; TITLE OF INVENTION: Genotype
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ronald J. Baron, Esq.
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,570A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/337,268
; FILING DATE: 11-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 454-3 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 146 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: no
US-08-484-570A-54

alignment_scores:
Quality: 72.00 Length: 37
Ratio: 2.880 Gaps: 0
Percent Similarity: 67.568 Percent Identity: 43.243

alignment_block:
US-09-647-780A-4 x US-08-484-570A-54
Align seg 1/1 to: US-08-484-570A-54 from: 1 to: 146

78 LysAlaTyrLeuLysTrpMetAlaGluGlyClyLysAspGlnGlnLeuPr 94
23 CAGCAATACAGCAGGAGCTGTACGCCACCATGGGAGACTGTCCTGCC 72
94 oGlyLeuAspLeuThrHisGluGlnLeuPhePheIleAsnTyrAlaGlnV 111
111 altrpCysGly 114
123 TAGCAGCGGC 133
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 11:05:27 ; Search time 30.1 seconds
(without alignments)
428.058 Million cell updates/sec

Title: US-09-647-780A-4
Perfect score: 665
Sequence: 1 GHEITHGDDNGRNFDPKNGN.....DLTHQQLFFINVAQVWGCK 116

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	651	97.9	691	22	AA83840 Amino acid sequenc
2	651	97.9	702	22	AA83844 Amino acid sequenc
3	651	97.9	753	22	AA83842 Amino acid sequenc
4	651	97.9	753	22	AA860562 Human neprilysin-1
5	651	97.9	762	22	AA860563 Human neprilysin-1
6	651	97.9	770	21	AA808131 A human neutral en
7	651	97.9	779	22	AA83841 Amino acid sequenc
8	651	97.9	779	22	AA860561 Human neprilysin-1
9	581	87.4	765	21	AA808130 A murine neutral e
10	578	86.9	774	20	AA44177 Rat membrane metal
11	471	70.8	750	9	AA82940 Atrial natriuretic

12	469	70.5	743	22	AAE07916 Human neutral endo
13	469	70.5	743	22	AAE07952 Human neutral endo
14	469	70.5	743	22	AA85104 Human enkephalinas
15	469	70.5	743	22	AA85115 Human enkephalinas
16	469	70.5	748	9	AA80969 Enkephalinase (hum
17	469	70.5	748	15	AA852706 Human enkephalinas
18	469	70.5	750	10	AA80393 Human common acute
19	465	69.9	255	13	AA82040 Short-form rat enk
20	465	69.9	750	9	AA82867 Enkephalinase (rat
21	465	69.9	750	15	AA82707 Rat enkephalinase.
22	397.5	59.8	661	22	AB861898 Drosophila melanog
23	374.5	56.3	189	16	AA876234 Bovine endothelin
24	374.5	56.3	708	16	AA876236 Bovine foetal hear
25	374.5	56.3	754	16	AA876226 Bovine endothelin
26	370.5	55.7	758	19	AA838369 Bovine endothelin
27	367.5	55.3	703	16	AA876237 Human placental en
28	367.5	55.3	753	16	AA876227 Endothelin convert
29	367.5	55.3	758	17	AA876227 Human endothelin c
30	367.5	55.3	758	17	AA876227 Human endothelin c
31	367.5	55.3	758	17	AA876227 Human endothelin c
32	353	53.1	749	19	AA853457 Amino acid sequenc
33	353	53.1	749	21	AA853457 Human PHEX. Homo
34	353	53.1	749	21	AA853457 Amino acid sequenc
35	343.5	51.7	883	22	AA867845 Polypeptide homolo
36	339.5	51.1	736	20	AA841774 Human PRO403 prote
37	339.5	51.1	736	21	AA844330 Human PRO403 prote
38	339.5	51.1	736	22	AA829233 Human PRO polypept
39	339.5	51.1	787	19	AA837750 Amino acid sequenc
40	338.5	50.9	763	22	AB864640 Drosophila melanog
41	336.5	50.6	567	16	AA876235 Bovine foetal hear
42	336	50.5	121	22	AA882596 Human secreted pro
43	323	48.6	726	22	AB861546 Drosophila melanog
44	320	48.1	112	22	AA882599 Human secreted pro
45	320	48.1	129	22	AA875602 Human colon cancer

ALIGNMENTS

RESULT 1
AA83840
ID AA83840 standard; Protein; 691 AA.
XX
XX
AA83840;
XX
XX
DT 23-JUL-2001 (first entry)
XX
XX
DE Amino acid sequence of a human metalloprotease enzyme IGS5.
XX
KW Metalloprotease; IGS5; infection; pain; cancer; diabetes; obesity; -
KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;
KW hypotension; hypertension; urinary retention; osteoporosis;
KW angina pectoris; myocardial infarction; stroke; ulcer; allergy;
KW benign prostatic hypertrophy; migraine; psychotic disorder;
KW neurological disorder; autism; multiple sclerosis; Alzheimer's disease;
KW neurodegenerative disease; sleep disorder; kidney disease;
KW cardiovascular disease; arteriosclerosis; cerebrovasospasm;
KW subarachnoid hemorrhage; cerebral ischemia; cerebral infarction;
KW peripheral vascular disease; Raynaud's disease; motility disorder;
KW gastrointestinal disorder; gastric emptying; gastroparesis; diarrhoea;
KW inflammation; chemotherapy induced injury; tumour invasion;
KW immune disorder; arthritis; endotoxin shock; sepsis; diabetes mellitus;
KW severe mental retardation; dyskinesia; Huntington's disease;
KW Gilles de la Tourette's syndrome.
OS Homo sapiens.
PN WO200136610-A1.
XX
XX
PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-EPI1532.
XX

QY 1 GHEITHGFDNCRNFKNGNMMDWSNFSTQHFRQSECMYQYGNYSWDLADEQNVNGF 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 535 gheithgfdngnfrndkngnmmdwsnfstqhfqseqcmlyqygnyswldadeqnvngf 594
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 NTLGENTADNGGVROQAYKAYLKWMAEGGKDQQLPGLDLTHTHEQLFFINVAQVWCG 114
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 595 ntlgeniadnggvrgaykaylkwmaeggdqqlpdlthcheqlffinyaqvwcg 648

RESULT 3

AAB83842
 ID AAB83842 standard; Protein: 753 AA.

XX AC AAB83842;

XX DT 23-JUL-2001 (first entry)

XX DE Amino acid sequence of a human metalloprotease enzyme IGS5.

XX KW Metalloprotease; IGS5; infection; pain; cancer; diabetes; obesity;
 KW anorexia; bulimia; asthma; Parkinson's disease; acute heart failure;
 KW hypotension; hypertension; urinary retention; osteoporosis;
 KW angina pectoris; myocardial infarction; stroke; ulcer; allergy;
 KW benign prostatic hypertrophy; migraine; psychotic disorder;
 KW neurological disorder; autism; multiple sclerosis; Alzheimer's disease;
 KW neurodegenerative disease; sleep disorder; epilepsy; kidney disease;
 KW cardiovascular disease; arteriosclerosis; cerebrovasospasm;
 KW subarachnoid hemorrhage; cerebral ischemia; cerebral infarction;
 KW peripheral vascular disease; Raynaud's disease; motility disorder;
 KW gastrointestinal disorder; gastric emptying; gastroparesis; diarrhoea;
 KW inflammation; chemotherapy induced injury; tumour invasion;
 KW immune disorder; arthritis; endotoxin shock; sepsis; diabetes mellitus;
 KW severe mental retardation; dyskinesia; Huntington's disease;
 KW Gilles de la Tourette's syndrome.

XX OS Homo sapiens.

XX PN WO2001136610-A1.

XX PD 25-MAY-2001.

XX PF 17-NOV-2000; 2000WO-EP11532.

XX PR 19-NOV-1999; 99EP-0203862.

XX PR 19-NOV-1999; 99NF-1013616.

XX PR 31-MAY-2000; 2000EP-0201937.

XX PR 31-MAY-2000; 2000NL-1015356.

XX PA (SOLV) SOLVAY PHARM BV.

XX PI Deleersnijder W, Wieggers R, Weske M;

XX DR WPI; 2001-343815/36.

XX DR N-PSDB; AAF89739.

XX PT New IGS5 polypeptides useful for treating infections, pain, cancer,
 PT diabetes, obesity, anorexia, bulimia, asthma, schizophrenia,
 PT hypertension, urinary retention and Parkinson's disease

XX PS Claim 3; Page 9-10; 115pp; English.

XX CC The present sequence represents a human metalloprotease enzyme designated
 CC IGS5. IGS5 polynucleotides and polypeptides are useful for treating
 CC infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma,
 CC Parkinson's disease, acute heart failure, hypotension, hypertension,
 CC urinary retention, osteoporosis, angina pectoris, myocardial infarction,
 CC stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,
 CC psychotic and neurological disorders, autism, multiple sclerosis,
 CC Alzheimer's disease, and other neurodegenerative diseases, sleep
 CC disorder, epilepsy, cardiovascular disease, arteriosclerosis,
 CC cerebrovasospasm, subarachnoid hemorrhage, cerebral ischemia, cerebral
 CC infarction, peripheral vascular disease, Raynaud's disease, kidney
 CC diseases, gastrointestinal disorders, motility disorders and conditions

CC of delayed gastric emptying, post-operative or diabetic gastroparesis,
 CC diarrhoea, inflammations, chemotherapy induced injury, tumour invasion,
 CC immune disorders, arthritis, endotoxin shock, sepsis, complications of
 CC diabetes mellitus, and severe mental retardation and dyskinesias, such
 CC as Huntington's disease or Gilles de la Tourette's syndrome.

XX SQ Sequence 753 AA;

Query Match 97.9%; Score 651; DB 22; Length 753;

Best Local Similarity 100.0%; Pred. No. 6,7e-64;

Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHEITHGFDNCRNFKNGNMMDWSNFSTQHFRQSECMYQYGNYSWDLADEQNVNGF 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 586 gheithgfdngnfrndkngnmmdwsnfstqhfqseqcmlyqygnyswldadeqnvngf 645
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 NTLGENTADNGGVROQAYKAYLKWMAEGGKDQQLPGLDLTHTHEQLFFINVAQVWCG 114
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 646 ntlgeniadnggvrgaykaylkwmaeggdqqlpdlthcheqlffinyaqvwcg 699
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 4

AAB60562

ID AAB60562 standard; Protein: 753 AA.

XX AC AAB60562;

XX DT 27-APR-2001 (first entry)

XX DE Human neprilysin-like membrane metalloproteinase SNEPB.

XX KW Human; SNEPB; neprilysin-like membrane metalloproteinase;
 KW splice variant; alternative splicing; zinc endopeptidase family;
 KW neuropeptide; peptide hormone; processing; metabolism; vaccine;
 KW drug screening; renal insufficiency; hepatic disease; ischaemia; pain;
 KW hypertension; cancer; inflammation; cardiovascular disease;
 KW neuronal disease; pancreatic disease; prostatic disease;
 KW respiratory disease; gene therapy; nephrotropic; vasotropic; analgesic;
 KW cerebroprotective; hypotensive; cytostatic; antinflammatory;
 KW cardiovascular; hepatotropic.

XX OS Homo sapiens.

XX PN EP1069188-A1.

XX PD 17-JAN-2001.

XX PF 15-JUL-1999; 99EP-0401767.

XX PR 15-JUL-1999; 99EP-0401767.

XX PA (SNFI) SANOFI-SYNTHELABO.

XX PI Jagerschmidt A, Agnel M, Culouscou J;

XX DR WPI; 2001-212582/22.

XX DR N-PSDB; AAF59660.

XX PT New membrane-associated metalloproteinase SNEPB, SNEPB and SNEPC
 PT polypeptides and polynucleotides, useful for treating e.g. acute and
 PT chronic renal insufficiency, pain, stroke, cancer, inflammation, renal
 PT and hepatic ischemia

XX PS Claim 11; Page 33-35; 72pp; English.

XX CC The invention relates to the human SNEPB, SNEPB and SNEPC proteins, and
 CC the cDNAs encoding them. SNEPB, SNEPB and SNEPC are neprilysin-like
 CC membrane metalloproteinases and are the products of alternative splicing.
 CC The substrate(s) for the SNEP proteins are not as yet known, although
 CC the neprilysin family of zinc endopeptidases play key roles in the
 CC processing and/or metabolism of neurotrophic and peptide hormones. SNEP
 CC nucleotides may be used as hybridisation probes for cDNA and genomic

CC DNA: to isolate full-length cDNAs and genomic clones encoding SNEPa,
 CC SNEPb or SNEPc; to isolate cDNA and genomic clones of SNEP homologues;
 CC as research reagents and material for the discovery of treatments and
 CC diagnostics for animal and human diseases; and for chromosome
 CC identification. The SNEP proteins may be used as immunogens to
 CC produce antibodies immunospecific for SNEPa, SNEPb or SNEPc. Such
 CC antibodies are used to isolate or identify clones expressing the
 CC protein, or to purify the proteins by affinity chromatography.
 CC SNEP proteins may also be used in screening for compounds which modulate
 CC SNEP endopeptidase activity, and to assess enzymatic cleavage of small
 CC molecule substrates in cells, cell-free preparations, chemical libraries
 CC and product mixtures. The SNEP proteins (as vaccine compositions),
 CC SNEP nucleotides, and SNEP activators or inhibitors may be used
 CC to treat acute and chronic renal insufficiency, renal and hepatic
 CC ischaemia, pain, stroke, hypertensive disease, cancer, inflammation, as
 CC well as cardiovascular, neuronal, pancreatic, prostatic, renal,
 CC respiratory or hepatic diseases. They may also be used in modulating
 CC peptide activation and/or degradation in the brain or kidney or in
 CC another organ, or to diagnose or treat any disorder related to abnormal
 CC expression of SNEPa, SNEPb or SNEPc. The present sequence represents
 CC SNEPb.
 XX
 XX Sequence 753 AA;

Query Match 97.9%; Score 651; DB 22; Length 753;
 Best Local Similarity 100.0%; Pred. No. 6.7e-64;
 Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHEITHGFDGNGRFDKNGNMMDWSNFSSTQHFRQSECMYQYGNYSWDLADEQNVNGF 60
 |||||
 Db 586 gheithgfdgngfrfdkngnmmdwsnfstqhrqsecmlyqygnyswldadeqnvngf 645
 |||||
 Qy 61 NTLGENIADNGGVRQAYKAYLKWMAEGKGDQOLPGLDLTHEQLFFINYAQVWCG 114
 |||||
 Db 646 ntlgeniadnggvrqaykaylkwmaegkgdqlpdlthelqffinyaqvwcg 699

RESULT 5
 AAB60563
 ID AAB60563 standard; Protein: 762 AA.
 XX
 AC AAB60563;
 XX
 DT 27-APR-2001 (first entry)
 XX
 DE Human neprilysin-like membrane metalloproteinase SNEPc.
 XX
 KW Human: SNEPc; neprilysin-like membrane metalloproteinase;
 KW splice variant; alternative splicing; zinc endopeptidase family;
 KW neuropeptide; peptide hormone; processing; metabolism; vaccine;
 KW drug screening; renal insufficiency; hepatic disease; ischaemia; pain;
 KW hypertension; cancer; inflammation; cardiovascular disease;
 KW neuronal disease; pancreatic disease; prostatic disease;
 KW respiratory disease; gene therapy; nephrotropic; vasotropic; analgesic;
 KW cerebroprotective; hypotensive; cytostatic; antiinflammatory;
 KW cardiovascular; hepatotropic.
 XX
 OS Homo sapiens.
 XX
 PN EP1069188-A1.
 XX
 PD 17-JAN-2001.
 XX
 PF 15-JUL-1999; 99EP-0401767.
 XX
 PR 15-JUL-1999; 99EP-0401767.
 XX
 PA (SNFI) SANOFI-SYNTHELABO.
 XX
 PI Jagerschmidt A, Agnel M, Culouscou J;
 XX
 DR WPI; 2001-212582/22.

DR N-PSDB; AAF59661.
 XX
 PT New membrane-associated metalloproteinase SNEPa, SNEPb and SNEPc
 PT polypeptides and polynucleotides, useful for treating e.g. acute and
 PT chronic renal insufficiency, pain, stroke, cancer, inflammation, renal
 PT and hepatic ischaemia.
 XX
 PS Claim 11; Page 38-40; 72pp; English.
 XX
 CC The invention relates to the human SNEPa, SNEPb and SNEPc proteins, and
 CC the cDNAs encoding them. SNEPa, SNEPb and SNEPc are neprilysin-like
 CC membrane metalloproteinases and are the products of alternative splicing.
 CC The substrate(s) for the SNEP proteins are not as yet known, although
 CC the neprilysin family of zinc endopeptidases play key roles in the
 CC processing and/or metabolism of neuropeptides and peptide hormones. SNEP
 CC nucleotides may be used as hybridisation probes for cDNA and genomic
 CC DNA; to isolate full-length cDNAs and genomic clones encoding SNEPa,
 CC SNEPb or SNEPc; to isolate cDNA and genomic clones of SNEP homologues;
 CC as research reagents and material for the discovery of treatments and
 CC diagnostics for animal and human diseases; and for chromosome
 CC identification. The SNEP proteins may be used as immunogens to
 CC produce antibodies immunospecific for SNEPa, SNEPb or SNEPc. Such
 CC antibodies are used to isolate or identify clones expressing the
 CC protein, or to purify the proteins by affinity chromatography.
 CC SNEP proteins may also be used in screening for compounds which modulate
 CC SNEP endopeptidase activity, and to assess enzymatic cleavage of small
 CC molecule substrates in cells, cell-free preparations, chemical libraries
 CC and product mixtures. The SNEP proteins (as vaccine compositions),
 CC SNEP nucleotides, and SNEP activators or inhibitors may be used
 CC to treat acute and chronic renal insufficiency, renal and hepatic
 CC ischaemia, pain, stroke, hypertensive disease, cancer, inflammation, as
 CC well as cardiovascular, neuronal, pancreatic, prostatic, renal,
 CC respiratory or hepatic diseases. They may also be used in modulating
 CC peptide activation and/or degradation in the brain or kidney or in
 CC another organ, or to diagnose or treat any disorder related to abnormal
 CC expression of SNEPa, SNEPb or SNEPc. The present sequence represents
 CC SNEPc.
 XX
 XX Sequence 762 AA;

Query Match 97.9%; Score 651; DB 22; Length 762;
 Best Local Similarity 100.0%; Pred. No. 6.8e-64;
 Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHEITHGFDGNGRFDKNGNMMDWSNFSSTQHFRQSECMYQYGNYSWDLADEQNVNGF 60
 |||||
 Db 595 gheithgfdgngfrfdkngnmmdwsnfstqhrqsecmlyqygnyswldadeqnvngf 654
 |||||
 Qy 61 NTLGENIADNGGVRQAYKAYLKWMAEGKGDQOLPGLDLTHEQLFFINYAQVWCG 114
 |||||
 Db 655 ntlgeniadnggvrqaykaylkwmaegkgdqlpdlthelqffinyaqvwcg 708
 |||||

RESULT 6
 AAB08131
 ID AAB08131 standard; Protein: 770 AA.
 XX
 AC AAB08131;
 XX
 DT 04-DEC-2000 (first entry)
 XX
 DE A human neutral endopeptidase metalloproteinase-like enzyme NL-2.
 XX
 KW Neprilysin; neutral endopeptidase metalloproteinase-like enzyme;
 KW NEP-like enzyme; protein production; protein secretion;
 KW neurological disease; Alzheimer's disease; pain; psychiatric disorder;
 KW fertility; bone disease; abnormal phosphate metabolism.
 XX
 OS Homo sapiens.
 XX
 PN WO2000/47750-A2.
 XX

RESULT 10

AA44177
ID AAY44177 standard; Protein; 774 AA.

XX AC
XX AAY44177;

DT 01-FEB-2000 (first entry)

DE Rat membrane metalloprotease NEPII protein.

XX Rat; membrane metalloprotease; neprilysine II; NEPII; inactivation;
KW neuron; hormone; peptide messenger; inhibitor; detection; disorder;
KW cardiovascular disease; neurodegenerative disease; growth disorder;
KW hypothalamic-hypophyseal axis; endocrine disorder.

XX Rattus rattus.

OS FR2777291-A1.

PN 15-OCT-1999.

PF 08-APR-1998; 98FR-0004389.

PR 08-APR-1998; 98FR-0004389.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PI Oulmet T, Gros C, Haret C, Bonhomme MC, Facchinetti P;
PI Schwartz JC;

DR WPI; 1999-593429/51.

DR N-PSDB; AAZ28810.

XX New membrane metalloprotease NEPII, involved in proteolysis of
PT neuronal and hormonal peptides, used to screen for inhibitors,
PT potentially useful for treating e.g. cardiovascular disease -

PS Claim 1; Page 12-16; 29pp; French.

XX This sequence represents the rat membrane metalloprotease designated
CC neprilysine II (NEPII), which is involved in (in)activation of neuronal
CC and hormonal peptide messengers. NEPII is used to screen for specific
CC substrates (used to detect NEPII in cells and tissues) or inhibitors,
CC which can also be used to detect NEPII or for treatment of disorders
CC related to peptidergic signalling in which NEPII is involved, e.g.
CC cardiovascular or neurodegenerative diseases; growth disorders of
CC endocrine origin; disturbances of the hypothalamic-hypophyseal axis
CC or endocrine disorders.

SQ Sequence 774 AA;

Query Match 86.9%; Score 578; DB 20; Length 774;

Best Local Similarity 85.1%; Pred. No. 9.7e-56;

Matches 97; Conservative 13; Mismatches 4; Indels 0; Gaps 0;

OY 1 GHEITHGFDNGRNFKNMMDWNSFSTQHFREQSECMYQYGNYSWDLADQNNGVF 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 607 gheithgfdngnrfknmmdwnsfstarfrqscmlygysnfweladnqnvgf 666

OY 61 NTGENTADNGGVRAQYKAYLKWMAEGGKDOQLPGLDLTHEQLFFINVAQVWCG 114
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 667 stlgeniadnggvraqyawlqlwlaeggrdrpnlntyaqlffinyaqvwcg 720

RESULT 11

AAP82940

ID AAP82940 standard; protein; 750 AA.

XX AC

XX AAP82940;

DT 23-NOV-1990 (first entry)

XX Atrial natriuretic polypeptide binding polypeptide.

DE Diuretic; atrium cardis; atrial natriuretic peptide binding;
KW hypotensive action.

XX JP63079598-A.

PN 09-APR-1988.

XX 22-SEP-1986; 86JP-0222192.

XX 22-SEP-1986; 86JP-0222192.

XX (SUNR) SUNTORY LTD.

XX WPI; 1988-137132/20.

XX N-PSDB; AAN81690.

XX Novel polypeptide, with diuretic action - is obtd. from atrium cardis
PT of mammals showing specific bond to atrial natriuretic polypeptide
PT and gene codinh it.

XX Disclosure; ; p; Japanese.

XX Peptide specifically binds to atrial natriuretic polypeptide (ANP)

CC and has diuretic (partic. natriuretic) and hypotensive action.

CC See also AAN81690-93 and AAP81282-P81309.

XX Sequence 750 AA;

Query Match 70.8%; Score 471; DB 9; Length 750;

Best Local Similarity 66.7%; Pred. No. 8.2e-44;

Matches 76; Conservative 24; Mismatches 14; Indels 0; Gaps 0;

OY 1 GHEITHGFDNGRNFKNMMDWNSFSTQHFREQSECMYQYGNYSWDLADQNNGVF 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 583 gheithgfdngnrfknmmdwntqgsannfkqsgcmvyqygnfswdlaggghlngi 642

OY 61 NTGENTADNGGVRAQYKAYLKWMAEGGKDOQLPGLDLTHEQLFFINVAQVWCG 114
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 643 ntigeniadnggvraqyayrayqnyvkngeekllpdlndhklfflnfaqvwcg 696

RESULT 12

AAE07916

ID AAE07916 standard; Protein; 743 AA.

XX AC

XX AAE07916;

DT 01-NOV-2001 (first entry)

DE Human neutral endopeptidase (NEP) protein.

XX Human; female sexual dysfunction; FSD; female sexual arousal dysfunction;
KW FSAD; neuropeptide Y; NPY; gynaecological; neutral endopeptidase; NEP;
KW enkephalinase; vasoactive intestinal peptide; VIP.

XX Homo sapiens.

XX EP1097718-A1.

PD 09-MAY-2001.

XX 03-NOV-2000; 2000EP-0309720.

XX 08-NOV-1999; 99GB-0026437.

XX 18-FEB-2000; 2000GB-0004021.

XX 26-MAY-2000; 2000GB-0013001.

XX 05-JUL-2000; 2000GB-0016563.

XX 12-JUL-2000; 2000GB-0017141.

PR 26-MAY-2000; 2000GB-0013001.
 PR 05-JUL-2000; 2000GB-0016563.
 PR 12-JUL-2000; 2000GB-0017141.
 XX
 PA (PFIZ) PFIZER LTD.
 PA (PFIZ) PFIZER INC.
 XX
 PI Maw GN, Wayman CP;
 XX
 XX
 DR WPI: 2001-383216/41.
 DR N-PSDB: AAF84073.
 XX

XX Novel pharmaceutical composition for treating female sexual
 PT dysfunction, preferably female sexual arousal disorder, comprising an
 PT agent capable of potentiating CAMP in the sexual genitalia of the
 PT female
 XX

PS Disclosure; Page 95; 132pp; English.

XX The invention provides a pharmaceutical composition for treating female
 CC sexual dysfunction (FSD), preferably female sexual arousal disorder
 CC (FSAD), or enhancing female genital (e.g. vaginal or clitoral) blood flow
 CC in a female. The composition comprises an agent capable of potentiating
 CC CAMP in the sexual genitalia of the female. The agent is an inhibitor
 CC (1:PDE) of phosphodiesterase (PDE), which hydrolyzes CAMP (optionally
 CC cGMP). The agent is useful in the manufacture of a medicament for the
 CC treatment of FSD, preferably FSAD, by potentiating CAMP in the sexual
 CC genitalia. It is also useful in the manufacture of a medicament for
 CC enhancing female genital (e.g., vaginal or clitoral) blood flow. The
 CC present sequence represents a human enkephalinase, a neutral
 CC endopeptidase (NEP) that can be used as an additional PCAMP target.
 XX

SQ Sequence 743 AA;

Query Match 70.5%; Score 469; DB 22; Length 743;
 Best Local Similarity 66.7%; Pred. No. 1.3e-43;
 Matches 76; Conservative 24; Mismatches 14; Indels 0; Gaps 0;

Oy 1 GHEITHGFDNDRNFDKNGNMOWNSFSTQHFRQSECMYQYGNYSWDLADQNVNWP 60
 Db 576 gheithgfdnfrnkdgdlvdtwtqgsanfkqsgcmvvyggnfswdlaggghlngi 635
 Oy 61 NTLGENIADNGVROAYKAYLKWMAEGGKDOQLPCLDLTHTHEOLFFINYAQVWCG 114
 Db 636 ntlgeniadngglgqayrayqnyikngeekilpdlidnkhqlflnfaqvwcg 689

RESULT 15

ID AAB85115 standard; Protein: 743 AA.

XX AAB85115;

XX 22-AUG-2001 (first-entry)

XX Human enkephalinase (EC 3.4.24.11).

XX Female sexual dysfunction; FSD; female sexual arousal disorder; FSAD;
 KW genital; vaginal; clitoral; blood flow; CAMP; phosphodiesterase; PDE;
 KW cGMP; medicament; human; enkephalinase; neutral endopeptidase; NEP.

OS Homo sapiens.

XX EP1097707-A1.

XX 09-MAY-2001.

XX 03-NOV-2000; 2000EP-0309719.

XX 08-NOV-1999; 99GB-0026437.

PR 18-FEB-2000; 2000GB-0004021.

PR 26-MAY-2000; 2000GB-0013001.

PR 05-JUL-2000; 2000GB-0016563.
 PR 12-JUL-2000; 2000GB-0017141.

XX (PFIZ) PFIZER LTD.
 XX (PFIZ) PFIZER INC.

XX Maw GN, Wayman CP;
 XX

XX WPI: 2001-383217/41.
 XX N-PSDB: AAF84083.

XX Novel pharmaceutical composition for treating female sexual
 PT dysfunction, preferably female sexual arousal disorder, comprising an
 PT agent capable of potentiating CAMP in the sexual genitalia of the
 PT female
 XX

PS Disclosure; Page 98; 135pp; English.

XX The invention provides a pharmaceutical composition for treating female
 CC sexual dysfunction (FSD), preferably female sexual arousal disorder
 CC (FSAD), or enhancing female genital (e.g. vaginal or clitoral) blood flow
 CC in a female. The composition comprises an agent capable of potentiating
 CC CAMP in the sexual genitalia of the female. The agent is an inhibitor
 CC (1:PDE) of phosphodiesterase (PDE), which hydrolyzes CAMP (optionally
 CC cGMP). The agent is useful in the manufacture of a medicament for the
 CC treatment of FSD, preferably FSAD, by potentiating CAMP in the sexual
 CC genitalia. It is also useful in the manufacture of a medicament for
 CC enhancing female genital (e.g., vaginal or clitoral) blood flow. The
 CC present sequence represents a human enkephalinase, a neutral
 CC endopeptidase (NEP) that can be used as an additional PCAMP target.
 XX

SQ Sequence 743 AA;

Query Match 70.5%; Score 469; DB 22; Length 743;
 Best Local Similarity 66.7%; Pred. No. 1.3e-43;
 Matches 76; Conservative 24; Mismatches 14; Indels 0; Gaps 0;

Oy 1 GHEITHGFDNDRNFDKNGNMOWNSFSTQHFRQSECMYQYGNYSWDLADQNVNWP 60
 Db 576 gheithgfdnfrnkdgdlvdtwtqgsanfkqsgcmvvyggnfswdlaggghlngi 635
 Oy 61 NTLGENIADNGVROAYKAYLKWMAEGGKDOQLPCLDLTHTHEOLFFINYAQVWCG 114
 Db 636 ntlgeniadngglgqayrayqnyikngeekilpdlidnkhqlflnfaqvwcg 689

Search completed: October 3, 2002, 11:07:38

Job time: 131 sec

Strd	Orig	zScore	EScore	Len	Documentation
9b_est2:BW459254	+ 493.00	1026.12	6.6e-48	1144	! BM459254 AGNCOURT_6415655 NIH
9b_est1:AJ396658	+ 455.00	1000.22	1.8e-46	667	! AJ396658 AGJ396658 dkf42426 Gallu
9b_est1:AW465461	+ 449.00	987.71	9.1e-46	610	! AW465461 BP230019B10E1l Soares
9b_est1:AI125432	+ 447.00	959.51	3.4e-44	705	! AI125432 BM44003.y1 Soares
9b_est1:BF968576	+ 428.00	936.74	6.3e-43	926	! BF968576 6022771182F1 NIH_MGC_84
9b_est1:AW940129	- 397.50	868.99	3.7e-39	899	! AW940129 GH03135.3prime GH_Dros
9b_est2:BI505516	- 389.50	853.93	2.6e-38	678	! BI505516 BB170022B10B07 Bee Bra
9b_est1:AI566727	- 386.00	850.32	4.1e-38	443	! AI566727 tf57aB04.x1 NCI_CGAP_K1
9b_est1:AA450725	+ 385.00	847.32	6.0e-38	479	! AA450725 vn18b04.x1 Soares_mamm
9b_est1:AI1913156	- 379.00	834.37	1.2e-37	458	! AI1913156 w44c01c0.x1 NCI_CGAP_K1
9b_est2:BE245519	+ 375.00	823.64	3.1e-36	550	! BE245519 TC8AP1D3248 Pediatric
9b_est2:BI692693	- 373.50	820.55	1.9e-36	536	! BI692693 603343357F1 NCI_CGAP_M
9b_est2:BE245100	+ 373.00	819.20	2.2e-36	549	! BE245100 TC8AP1D2555 Pediatric
9b_est2:W89492	+ 372.50	818.99	2.3e-36	501	! W89492 mf70d09.y1 Soares mouse
9b_est2:BI869244	+ 367.50	801.33	2.2e-35	968	! BI869244 603396112F1 NIH_MGC_90
9b_est1:AA895479	+ 366.50	807.27	1.0e-35	423	! AA895479 v534a09.l1 Stratagene
9b_est2:BJ093321	+ 360.50	790.02	9.9e-35	626	! BJ093321 BJ093321 N1BB Mochii n
9b_est2:BJ096397	+ 360.50	789.63	9.8e-35	651	! BJ096397 BJ096397 N1BB Mochii n
9b_est2:W10214	+ 352.00	773.97	7.3e-34	466	! W10214 ms44c03.l1 Soares mouse
9b_est2:W78610	+ 351.50	772.46	8.9e-34	485	! W78610 me86a0e.l1 Soares mouse
9b_est1:AI671752	- 350.00	769.76	1.3e-33	454	! AI671752 wa05e11.x1 NCI_CGAP_K1
9b_est2:BF476098	+ 350.00	768.05	1.6e-33	540	! BF476098 hr77d03.x1 NCI_CGAP_K1
9b_est1:AI702551	- 346.00	760.37	4.2e-33	476	! AI702551 tz94h04.x1 NCI_CGAP_K1
9b_est2:BI360370	+ 343.50	752.38	1.2e-32	608	! BI360370 387247 MARC 2P1G Sus
9b_est2:BI3603165	+ 342.50	751.75	1.3e-32	517	! BI360316 fp86f11.y3 Zebrafish a
9b_est2:BF590024	- 341.00	751.16	1.4e-32	391	! BF590024 ns448908.l1 NCI_CGAP_K
9b_est1:AW086484	- 339.00	743.10	3.8e-32	563	! AW086484 CG474902.x1 NCI_CGAP_Ov
9b_est2:BE871194	+ 337.50	736.78	8.6e-32	761	! BE871194 601448772F1 NIH_MGC_65
9b_est2:BF044240	+ 337.00	738.14	7.2e-32	592	! BF044240 BP250002B30D9 Soares r
9b_est2:BF042751	+ 335.50	734.90	1.1e-31	585	! BF042751 603065563F3 NIH_MGC_11
9b_est2:BE535389	+ 333.50	728.78	2.4e-31	693	! BE535389 601058542F1 NIH_MGC_10
9b_est2:BF124139	+ 332.50	723.91	4.5e-31	905	! BF124139 r6760084F1 NCI_CGAP_M
9b_est1:AW828182	+ 325.00	712.04	2.1e-30	553	! AW828182 ra07e10.y1 Bird-Rao Me
9b_est2:BF663245	+ 318.50	692.86	2.4e-29	889	! BF663245 602144363F1 NIH_MGC_48
9b_est2:BG803768	+ 314.50	687.81	4.6e-29	600	! BG803768 0241-58 Mouse E14.5 re
9b_est2:BF986454	+ 314.50	687.81	4.6e-29	600	! BF986454 3171-80 Mouse E14.5 re
9b_est1:AA028052	+ 314.00	689.85	3.5e-29	436	! AA028052 z33d12.y1 Bird-Rao Me
9b_est1:AA026997	+ 313.00	689.59	3.7e-29	357	! AA026997 re99e01.y1 Soares_feta
9b_est1:AI874916	- 312.00	690.91	3.1e-29	245	! AI874916 ul27a03.x1 Sugano mouse
9b_est1:AV329959	+ 311.50	680.11	1.2e-28	665	! AV329959 AV329959 RIKEN full-l

1 GlyHisGluIleThrHisGlyPheAspAspAsnGlyArgAsnPheAspLy 17
20 GGTCAAGAAATCACACACGGTTTTGATGATCAACCGCAGAAATTTTCAATGA 369
17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
70 GAATGAGAGACCTCGTAGACTGGTGGACTGAAGAAATCTGCACGTAACTTTA 419
34 rgGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAsp 50

194 a 124 c 139 g 151 t 2 others

BASE COUNT
ORIGIN

This read has been verified (found to hit its original self in the correct orientation)

Seq primer: -40RP from Gibco
High quality sequence stop: 459.

FEATURES

Location/Qualifiers
1..705
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:313540"
/clone_lib="Soares mouse p3NMF19.5"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCCGCAATTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."

BASE COUNT 218 a 142 c 163 g 177 t
ORIGIN

alignment_scores:
Quality: 437.00 Length: 114
Ratio: 4.243 Gaps: 0
Percent Similarity: 90.351 Percent Identity: 63.158

alignment_block:

US-09-647-780A-4 x AI325325

Align seg 1/1 to: AI325325 from: 1 to: 705

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17
195 GGACATGAATCACACATGGATTGATGACAAATGGAAGAAATTTCAATA 244
17 sAsnGlyAsnMetMetAspTyrSerAsnPheSerThrGlnHisPheA 34
245 AGATGGAGACCTCGTGTACTGGTGGACGAGCAGTCGCAAAATTTTA 294
34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTyrPasp 50
295 AAGACCAATCCAGTGTATGGTATACCATGTAAGAACTTTTCTCTGTGAC 344
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsn 67
345 CTAGCAGGTGGACGACATCTCAATGGAATTAATACACTAGGAGAAACAT 394
67 eAlaAspAsnGlyValArgGlnAlaTyrLysAlaTyrLeuLysTyrPM 84
395 TCGTGTATATGAGGTATTGGCCAGCATACAGAGCCTATCAGAAATATG 444
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
445 TTAAGAAGAGTGTGAAGAAATTAATCTCCCTGACTTGACCTCAATCAC 494
101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTyrCysGly 114
495 AAACAACATATNTTCTGAACTGTGCCGAGTGTGTGTGGG 536

seq_name: gb_est2:BF968576

seq_documentation_block:

LOCUS BF968576 926 bp mRNA linear EST 22-JAN-2001
DEFINITION '602271182F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4359141 5',
mRNA sequence.
ACCESSION BF968576
VERSION BF968576.1 GI:12335791

alignment_scores:
Quality: 449.00 Length: 114
Ratio: 4.317 Gaps: 0
Percent Similarity: 91.228 Percent Identity: 63.158

alignment_block:

US-09-647-780A-4 x AW465461

Align seg 1/1 to: AW465461 from: 1 to: 610

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17
218 GGACAGAAATCACCCATGGCTTCGATGACATGGCAGAAATTTAAACA 267
17 sAsnGlyAsnMetMetAspTyrTrpSerAsnPheSerThrGlnHisPheA 34
268 GGATGGAGACCTTGTGACTGTGGACTCAACAGCTCTGCAAAATATTTTA 317
34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTyrPasp 50
318 AAGACCTATCCAGTGTGATGGTACCATGATGGAACTTCTCTGGGAC 367
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsn 67
368 CTACCAATGGACAGCATCTCAATGGAATTAATACATGGGAGAAACAT 417
67 eAlaAspAsnGlyValArgGlnAlaTyrLysAlaTyrLeuLysTyrPM 84
418 TGCATGATAATGGTGTATGGCCAGCATACAGAGCCATCAAAACTATG 467
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
468 TTAAGAAGAGTGTGAAGAAATTAATCTCTGGACTTGACCTTAATCAC 517
101 GluGlnLeuPhePheIleAspTyrAlaGlnValTyrCysGly 114
518 AGACAACTGTCTTCCGTAATTTGCCCANNTGTGGTGTGGA 559

seq_name: gb_est1:AI325325

seq_documentation_block:

LOCUS AI325325 705 bp mRNA linear EST 23-DEC-1998
DEFINITION ma44c03.y1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
IMAGE:313540 5' similar to gb:J03779 NEPRILYSIN (HUMAN);, mRNA
sequence.

ACCESSION AI325325

VERSION AI325325.1 GI:4059754

KEYWORDS EST

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 705)

AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MG1:204156

This read is a RESEQUENCE of a previously sequenced mouse clone

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 926)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLM9998 row: d column: 22
High quality sequence stop: 766.

FEATURES
source
1..926
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4359141"
/clone_lib="NIH_MGC_84"
/tissue_type="adrenal cortex carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 286 a 192 c 221 g 227 t
ORIGIN

alignment_scores
Quality: 428.00 Length: 112
Ratio: 4.076 Gaps: 0
Percent Similarity: 93.750 Percent Identity: 65.179

alignment_block
US-09-647-780A-4 x BF968576
Align seg 1/1 to: BF968576 from: 1 to: 926

3 Glu1e1rHisGlyPheAspAsnGlyArgAsnPheAspLysAsnG1 19
|||||
2 GAATCACCCTGGCTTCGATGACAAAGCGCAAACTTAAACAAGATGG 51
|||||

19 yAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheArgGluG 36
|||||
52 AGACCTCGTTGACTGGTGGAGTCAACAGTGTGCAAGTAACCTTAAAGGAGC 101
|||||

36 lnsrGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpAspLeuAla 52
|||||
102 AATCCCAAGTCATGGTATCAGTATGGAACATTTCTCGGAGCTGGCA 151
|||||

53 AspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnIleAlaAs 69
|||||
152 GGTGCAGCAGCACCCTTAATGGAATTAATACACTGGGAGAAACAT GCTGA 200
|||||

69 pAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpMetAlaG 86
|||||
201 TAATGGAGGTCTTGTCAAGCATACAGAGCCTATCAGAAATATATATAAA 250
|||||

86 luGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHisGluGln 102
|||||
251 AGAATGCGGAGAAATAATCTCTCGACTTGACCTAAATCAACAACAA 300
|||||

103 LeuPhePheIleAsnTyrAlaGlnValTrpCysGly 114
|||||

301 CTATTTTCTTGAACCTTTCACAGGTGTGGTGTGGA 336
seq_name: gb_est:AW940129
seq_documentation_block:
LOCUS AW940129.1 899 bp mRNA linear EST 23-APR-2001
DEFINITION GH03315.3prime GH Drosophila melanogaster head pot2 Drosophila melanogaster cDNA clone GH03315 3', mRNA sequence.
ACCESSION AW940129
VERSION AW940129.1 GI:8115575
KEYWORDS EST
SOURCE fruit fly
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 899)
AUTHORS Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G., Lewis, S. and Rubin, G.M.
TITLE BDGP/HMI Drosophila EST Project
JOURNAL Unpublished (2001)
COMMENT Other_ESTs: GH03315.5prime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd., Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or more T residues at the beginning of the sequence, this clone was polyadenylated. The resulting Poly-T sequence has been removed. hit genomic AE003436: Drosophila melanogaster genomic scaffold 142000013386054 section 20 of 35, complete sequence.: 02/14/2001
Plate: GH.33 row: B column: 3
High quality sequence stop: 819.

FEATURES
source
1..899
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="GH03315"
/clone_lib="GH Drosophila melanogaster head pot2"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"
/notes="Organ: head; Vector: pot2; Site: 1; EcoRI; Site: 2; XhoI; Sized fractionated cDNAs were directly ligated into pot2. Plasmid cDNA library."
BASE COUNT 213 a 223 c 240 g 221 t
ORIGIN

alignment_scores
Quality: 397.50 Length: 114
Ratio: 4.015 Gaps: 1
Percent Similarity: 86.842 Percent Identity: 59.649

alignment_block
US-09-647-780A-4 x AW940129/rev
Align seg 1/1 to reverse of: AW940129 from: 1 to: 899

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAspLys 17
|||||
830 GCCACGAGATTACCCAGGATTCGACGACAGGCGCAGCGCTGCACAA 781
|||||

17 sAsnGlyAsnMetMetAspTrpTrpSerAsnPheSerThrGlnHisPheA 34
|||||
780 GGAGGCAACATGATGCAGTGTGGACAATGCCACCATCGAGGCGCTTC 731
|||||

34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrpasp 50
|||||
730 CGCAACGCGACGAGTGGCTCATCGATCAGTACTCGCTACAAGATTAC. 681
|||||

REFERENCE
1 (bases 1 to 479)
AUTHORS
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

1. The first group of students (Group A) was assigned to the traditional lecture method. They received a 45-minute lecture on the topic of "The Role of the Teacher in the Classroom." The lecture was delivered by the instructor, who used a power point presentation to illustrate key points. The students were then given a 15-minute quiz to assess their understanding of the material.

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae

BASE COUNT 184 a 110 c 124 g 131 t 1 others
ORIGIN

alignment_scores:
Quality: 375.00 Length: 100
Ratio: 4.032 Gaps: 0
Percent Similarity: 93.000 Percent Identity: 65.000

alignment_block:
US-09-647-780A-4 x BE245519

Align seg 1/1 to: BE245519 from: 1 to: 550

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17
|||||
250 GGACAGAAATCACCCTGCTTCGATGCAATGCGAGAACTTTAAACA 299
17 sAsnGlyAsnMetMetAspTrpSerAsnPheSerThrGlnHisPheA 34
|||||
300 AGATGGAGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 349
34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrp 50
|||||
350 AGGAGCAATCCAGTCGATGCTGATGATGATGATGATGATGATGAT 399
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAs 67
|||||
400 CTGGCAGGTGGACAGCACCCTTAATGGAATTAATACACTGGGAGAA 449
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLys 84
|||||
450 TGCTGATATAGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 499
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThr 100
|||||
500 TTAAGAAGATGGCGAAGAAATTA.CTTCTGGACTTGACCTAAATCAC 548

seq_name: gb_est2:B1692693

seq_documentation_block:
LOCUS B1692693 536 bp mRNA linear EST 18-SEP-2001
DEFINITION 60343357F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5370927 5',
mRNA sequence.

ACCESSION B1692693

VERSION B1692693.1 GI:15655322

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

1 (bases 1 to 536)

NTN-MGC <http://mgs.nci.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1A11944 Row: b Column: 16

High quality sequence stop: 536.

Location/Qualifiers

1..536

/organism="Mus musculus"

/strain="FVB/N-3"

/db_xref="taxon:10090"

/clone="IMAGE:5370927"

/clone_lib="NCI_CGAP_Mam2"

FEATURES

Source

/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 99 a 162 c 156 g 119 t
ORIGIN

alignment_scores:
Quality: 373.50 Length: 116
Ratio: 3.851 Gaps: 1
Percent Similarity: 83.621 Percent Identity: 54.310

alignment_block:
US-09-647-780A-4 x B1692693/rev

Align seg 1/1 to reverse of: B1692693 from: 1 to: 536

1 GlyHisGluIleThrHisGlyPheAspAsnGlyArgAsnPheAsp 17
|||||
419 GGCCACGAGCTGACTCAGCTTCGATGATCAAGTCGGGAGTACGACA 370
17 sAsnGlyAsnMetMetAspTrpSerAsnPheSerThrGlnHisPheA 34
|||||
369 GGATGGAACTTCGACCTCGTGGGAAGACTCGCTCCGTGGAGCGCTTCA 320
34 rgGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrp 50
|||||
319 ACCAGCAGCAGGAGTGCATGGTGCACATAACAGCAACTACAGC... 276
51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAs 67
|||||
275 ...GTGAACGGGGAGCCCGTGAACGGTGCACACACCCCTCGGGGAGA 229
67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLys 84
|||||
228 TCCCGACAACGGGGGACTCAAGCGGCCCTACCGGGCGTACCAGAACTGGG 179
84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThr 100
|||||
178 TAAAGAGAAGCAGGAGCTGAGCAGATGCTGCCACCCCTGGGTCTCAC 129
101 GluGlnLeuPhePheIleAsnTyrAlaGlnValTrpCysGlyCys 116
|||||
128 AACCAGCTCTTCTCTAGGATTGACAGAGTCTGTGTCTCTGTCTCGC 81

seq_name: gb_est2:BE245100

seq_documentation_block:

LOCUS BE245100 549 bp mRNA linear EST 03-OCT-2001

DEFINITION TCBAPID2535 Pediatric pre-B cell acute lymphoblastic leukemia

Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCBAP2535, mRNA.

Sequence.

ACCESSION BE245100

VERSION BE245100.1 GI:9096842

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 549)

AUTHORS Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman Jr., F.R., Muzny, D.

, Bouck, J., Gibbs, R.A. and Margolin, J.F.

Pediatric Leukemia cDNA Sequencing Project

Unpublished (2000)

CONTACT: Dr. Judith F. Margolin

Texas Children's Cancer Center and Human Genome Sequencing Center

at Baylor College of Medicine

1102 Bates, MC3-3320 Houston, TX 77030, USA

Tel: 832-824-4536

seq documentation block:

17 sAsnGlyAsnMetMetAspTrrpSerAsnPheserThrGlnHisPhea 34
 57 GGATGGACCTTCGACCCCTGGTGGAGAACTCGTCGGTGGAGCGGTCA 106
 34 rGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrrpAsp 50
 107 AGCAGCAGACGGAGTGCATGGTGCAGCAATACAGCAACTACAGC..... 150
 51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnI 67
 151 ..GTAACGGGAGCGCGGTGACAGCTCGACACACCCCTCGGGGAGAACAT 197
 67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
 198 TGCGGACACGGGGGACTCAAGCGGCGCTACCGGGGTACCAAGACTGGG 247
 84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
 248 TAAAGAAGAACGGAGCTGAGCAGACGCTGCCACCCCTGGGTCTCACCAGC 297
 101 GluGlnLeuPhePheLeuAsnTyrAlaGlnValTrrpCysGlyCysLys 116
 298 AACCAAGCTCTTCTCTAGGATTGCACAGGTCTGGTCTCTGCCGC 345

seq_name: gb_est2:BI869244

seq_documentation_block: 968 bp mRNA linear EST 11-OCT-2001
 LOCUS BI869244 603396112F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5406003.5
 DEFINITION mRNA sequence.

ACCESSION BI869244

VERSION BI869244.1 GI:16042917

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 968)

NIH-MGC <http://mgc.ncl.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM12035 row: h column: 04

High quality sequence stop: 839.

Location/Qualifiers

FEATURES

source

1..968

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5406003"

/tissue_type="adrenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: liver; Vector: pCMV-SPORT6; Site: 1; NotI;

Site: 2; Salt: Cloned unidirectionally; oligo-dT primed.

Average insert size 1.7 Kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 214 a 308 c 277 g 168 t

ORIGIN

alignment_scores:

Quality: 367.50

Ratio: 3.750

Length: 116

Gaps: 1

Percent Similarity: 84.483 Percent Identity: 53.448

alignment_block:

US-09-647-780A-4 x BI869244

Align seg 1/1 to: BI869244 from: 1 to: 968

1 GYHisGluIleThrHisGlyPheAspAsnGlyArgAsnPhesAspLy 17
 192 GGCCATGAGCTGACTCATGCTTTTGTATGATCAAGGACGGAGTATGACAA 241
 17 sAsnGlyAsnMetMetAspTrrpSerAsnPheserThrGlnHisPhea 34
 242 GGACGGGAACCTCCGGCCATGGTGAAGAACTCATCGTGGAGGCCTTCA 291
 34 rGluGlnSerGluCysMetIleTyrGlnTyrGlyAsnTyrSerTrrpAsp 50
 292 AGCGTCAGACCGGAGTGTGATGAGCAGTACAGCAACTACAGC..... 335
 51 LeuAlaAspGluGlnAsnValAsnGlyPheAsnThrLeuGlyGluAsnI 67
 336 ..GTGAACGGGAGCGGTGAAGGGGGGGGACACACCTCGGGGAGAACAT 382
 67 eAlaAspAsnGlyGlyValArgGlnAlaTyrLysAlaTyrLeuLysTrpM 84
 383 CGCGCACACGGGGGTCTCAAGCGGCGCTATCGGGCTTACCAGAACTGGG 432
 84 eAlaGluGlyGlyLysAspGlnGlnLeuProGlyLeuAspLeuThrHis 100
 433 TGAAGAAGAACGGGGGTGAGCAGCTCGCTCCCGCCCTGGGGCTCACCNA 482
 101 GluGlnLeuPhePheLeuAsnTyrAlaGlnValTrrpCysGlyCysLys 116
 483 AACCAAGCTCTTCTCTAGGATTGCACAGGTCTGGTCTCTGCCGC 530

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 11:05:27 ; Search time 11.88 Seconds
(without alignments)
378.070 Million cell updates/sec

Title: US-09-647-780A-4
Perfect score: 665
Sequence: 1 GHEITHGFDNCRNFKNGN.....DLTHEQLFFINTAQVWCGCK 116

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	469	70.5	749	1 NEP_HUMAN	P08473 homo sapien
2	469	70.5	749	1 NEP_RABIT	P08049 cryotolagus
3	468	70.4	749	1 NEP_MOUSE	O61391 mus musculus
4	465	69.9	749	1 NEP_RAT	P07861 rattus norv
5	374.5	56.3	754	1 ECE1_BOVIN	P42891 bos taurus
6	370.5	55.7	754	1 ECE1_CAVPO	P97739 cavia porce
7	370.5	55.7	762	1 ECE1_RAT	P42893 rattus norv
8	367.5	55.3	770	1 ECE1_HUMAN	P42892 homo sapien
9	353	53.1	749	1 PEX_HUMAN	P78562 homo sapien
10	351	52.8	749	1 PEX_MOUSE	P70669 mus musculus
11	343.5	51.7	765	1 ECE2_HUMAN	O60344 homo sapien
12	339.5	51.1	787	1 ECE2_BOVIN	Q10711 bos taurus
13	263	39.5	769	1 YCVL_CABEL	Q22523 caenorhabdi
14	166.5	25.0	647	1 PPO_LACHE	O52071 lactobacill
15	165	24.8	626	1 PPO_LACLA	Q07744 lactococcus
16	165	24.8	626	1 PPO_LACLC	Q09145 lactococcus
17	163.5	24.6	564	1 YSC6_STRGC	P42359 streptococc
18	113	17.0	732	1 KELL_HUMAN	P23276 homo sapien
19	78	11.7	543	1 PROA_LEGPN	P21347 legionella
20	75.5	11.4	256	1 YREC_SYNP2	P19737 synechococc
21	74	11.1	529	1 PROA_LEGLO	P55110 legionella
22	73.5	11.1	1078	1 GYRB_SYNP3	P77966 synechocyst
23	71.5	10.8	328	1 Y915_STRPY	Q9A064 streptococc
24	70	10.5	999	1 RLK5_ARATH	P47735 arabidopsis
25	68.5	10.3	730	1 GLGB_HAEIN	P45177 haemophilus
26	68	10.2	565	1 HEMA_TAHUR	P17002 influenza a
27	68	10.2	647	1 CMF_ECOLI	P33927 escherichia
28	68	10.2	731	1 BAF1_YEAST	P14164 saccharomyc
29	67	10.1	362	1 COTH_BACSU	O45535 bacillus su
30	67	10.1	608	1 VATI_BORBU	O51118 borrelia bu
31	66.5	10.0	328	1 Y095_STRPN	Q97160 streptococc
32	66.5	10.0	526	1 CLOS_CLOHI	P09870 clostridium
33	66.5	10.0	648	1 CMF_HAEIN	P45037 haemophilus

34	66	9.9	632	1 PARE_HAEIN	P43703 haemophilus
35	66	9.9	898	1 IF38_CABEL	O02328 caenorhabdi
36	65.5	9.8	224	1 MTGA_ACICA	O24849 acinetobact
37	65.5	9.8	847	1 HEX_VIBVU	O04786 vibrio vuln
38	65	9.8	265	1 SR21_SARPE	P24491 sarcophaga
39	65	9.8	500	1 ABFA_BACSU	P94531 bacillus su
40	65	9.8	942	1 HEX_ADEGL	P42671 avian adeno
41	64.5	9.7	627	1 TSD2_ABIGR	O24474 abies grand
42	64.5	9.7	1157	1 XNA_THESA	P36917 thermonaer
43	64.5	9.7	1241	1 TRK1_SACBA	P28569 saccharomyc
44	64.5	9.7	2813	1 VWF_CANFA	Q08295 canis famil
45	64	9.6	631	1 THIC_SALTY	Q01917 salmonella

ALIGNMENTS

RESULT 1

ID	NEP_HUMAN	STANDARD	PRT	749 AA
AC	P08473			
DT	01-AUG-1988	(Rel. 08, Created)		
DT	01-AUG-1988	(Rel. 08, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Neprilysin (EC 3.4.24.11) (Neutral endopeptidase) (NEP)			
DE	(Enkephalinase) (Common acute lymphocytic leukemia antigen) (CALLA)			
DE	(Neutral endopeptidase 24.11) (CD10)			
GN	MME OR EPN			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo			
OX	NCBI_TaxID=9606			
RN	[1]			
RP	SEQUENCE OF 2-749 FROM N.A.			
RC	TISSUE=Placenta			
RX	MEDLINE=88152222; PubMed=3162217			
RA	Malfroy B., Kuang W.-J., Seeburg P.H., Mason A.J., Schofield P.R.;			
RT	"Molecular cloning and amino acid sequence of human enkephalinase (neutral endopeptidase)."			
RL	FEBS Lett. 229:206-210(1988)			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney			
RX	MEDLINE=89010526; PubMed=2971756			
RA	Letarte M., Vera S., Tran R., Addis J.B.L., Onizuka R.J.,			
RT	Quackenbush E.J., Jongeneel C.V., McInnes R.R.;			
RL	"Common acute lymphocytic leukemia antigen is identical to neutral endopeptidase."			
RN	[3]			
RP	J. Exp. Med. 168:1247-1253(1988)			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	MEDLINE=88263038; PubMed=2968607			
RA	Shipp M.A., Richardson N.E., Sayre P.H., Brown N.R., Masteller E.L.,			
RT	Clayton L.K., Ritz J., Reinherz E.L.;			
RL	"Molecular cloning of the common acute lymphoblastic leukemia antigen (CALLA) identifies a type II integral membrane protein."			
RN	[5]			
RP	Proc. Natl. Acad. Sci. U.S.A. 85:4819-4823(1988)			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RC	MEDLINE=89386688; PubMed=2528730			
RA	D'Adamo L., Shipp M.A., Masteller E.L., Reinherz E.L.;			
RT	"Organization of the gene encoding common acute lymphoblastic leukemia antigen (neutral endopeptidase 24.11): multiple minlexons and separate 5' untranslated regions."			
RN	[7]			
RP	Proc. Natl. Acad. Sci. U.S.A. 86:7103-7107(1989)			
RN	[8]			
RP	ACTIVE SITE ASP-650			
RC	MEDLINE=94222095; PubMed=8168535			
RA	Le Moual H., Dion N., Roques B.P., Crine P., Boileau G.;			
RT	"Asp650 is crucial for catalytic activity of neutral endopeptidase 24-11."			
RN	[9]			
RP	Eur. J. Biochem. 221:475-480(1994)			

Db 642 NTLGENIADNGGIGQAYRAYQNYVKKNGEKKLLPGLDLNHHKQLFFLNFAQVWCG 695

	RESULT	4			
	NBP_RAT				
ID	NBP_RAT		STANDARD;	PRT;	749 AA..
AC	P07861:				
DT	01-AUG-1988	(Rel. 08,	Created)		
DT	01-AUG-1988	(Rel. 08,	Last sequence update)		
DT	16-OCT-2001	(Rel. 40,	Last annotation update)		
DE	Neprilysin	(EC 3.4.24.11)	(Neutral endopeptidase) (NEP)		
GN	(Enkephalinase).				
MFE					

OS *Rattus norvegicus* (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OOC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N. A., AND PARTIAL SEQUENCE.
 RX MEDLINE=87213218; PubMed=3555489;
 RA Malliroy, B., Schofield P.R., Kuang W.-J., Seeburg P.H., Mason A.J.,
 RA Henzel W.J.;
 RA "Molecular cloning and amino acid sequence of rat enkephalinase.";
 RL Biochem. Biophys. Res. Commun. 144:59-66(1987).

[2] RP ACTIVE SITE ARG-102.
RX MEDLINE=89197908; PubMed=2703483;
RA Mooman R.C., Jr., Jackson D., Slaughter C.A., Unnithan S., Chai Y.G.,
RA Boman C., Herish L.B.;
RT "Identification of the active-site arginine in rat neutral
RT endopeptidase 24.11 (enkephalinase) as arginine 102 and analysis of a
RT glutamine 102 mutant.";
J. Biol. Chem. 264:6151-6157(1989).
CC -!- FUNCTION: THERMOLYSIN-LIKE SPECIFICITY, BUT IS ALMOST CONFINED ON
CC ACTING ON POLYPEPTIDES OF UP TO 30 AMINO ACIDS. BIOLOGICALLY
CC IMPORTANT IN THE DESTRUCTION OF OPIOID PEPTIDES SUCH AS MET-
CC AND LEU-ENKEPHALINS BY CLEAVAGE OF A GLY-PHE BOND.
CC -!- COFACTOR: BINDS 1 ZINC ION.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M13 (ZINC
CC METALLOPROTEASE); ALSO KNOWN AS THE NEPRILYSIN SUPERFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

[illegible]

	650	650	PROTON DONOR (BY SIMILARITY).
FT ACT_SITE	650	650	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	144	144	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	210	210	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	284	284	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	310	310	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	324	324	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	627	627	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE	749 AA, 85663 MW;	89B0EC50A9016A47 CRC64;	(BY SIMILARITY).

Query Match	69.9%	Score 465;	DB 1;	Length 749;
Best Local Similarity	64.9%;	pred. No. 1.9e-39;		
Matches 74;	Conservative 26;	Mismatches 14;	Indels 0;	Gaps

[illegible]

RESULT	5	
ECE1_BOVIN		
ID	ECE1_BOVIN	STANDARD; PRT; 754 AA.

AC P42891;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Endothelin-converting enzyme 1 (EC 3.4.24.71) (ECE-1).
 GN EC1.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 CX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE=95104223; PubMed=7805846;
 RA Schmidt M., Kroeger B., Jacob E., Seulberger H., Subkowski T.,
 RA Otter R., Meyer T., Schmalzing G., Hillen H.;
 RT "Molecular characterization of human and bovine endothelin converting
 RT enzyme (ECE-1).";
 RL FEBS Lett. 356:238-243(1994).
 CC -1- FUNCTION: CONVERTS BIG ENDOTHELIN-1 TO ENDOTHELIN-1.
 CC -1- CATALYTIC ACTIVITY: Forms endothelin 1 by cleavage of the 21-
 CC Trp-1-Val-22 bond in the precursor.
 CC -1- ENZYME REGULATION: INHIBITED BY PHOSPHORAMIDON.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M13 (ZINC
 CC METALLOPROTEASE); ALSO KNOWN AS THE NEPRILYSIN SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on the use of this entry by non-profit institutions as long as its content is not modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/annotation> or send an email to license@sib-sb.ch).

CC	EMBL; 235306; CA84547.1; -	
DR	HSSP; P08473; IDMT.	
DR	MEROPS; M13:002; -	
DR	InterPro: IPR000718; Peptidase_M13.	
DR	InterPro: IPR000130; Zn_MTPeptidse.	
DR	Pfam; PF01431; Peptidase_M13; 1.	
DR	PROSITE; PRO00786; NEPRILYSIN.	
DR	PROSITE; PS00142; ZINC_PROTEASE; 1.	
KW	Hydrolase; Metalloprotease; Zinc; Glycoprotein; Transmembrane;	
KW	Signal-anchor.	
FT	DOMAIN 1 52 CYTOPLASMIC (POTENTIAL);	

Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
[7]
SEQUENCE OF 1-132 FROM N.A. (ISOFORM D), AND TISSUE SPECIFICITY.
MEDLINE-99421637; PubMed-10491078;
Valdenaire O., Lepailleur-Enouf D., Egidy G., Thouard A., Barret A.,
Vranckx R., Tougard C., Michel J.-B.,
RT "A fourth isoform of endothelin-converting enzyme (ECE-1) is generated
from an additional promoter.",
Eur. J. Biochem. 264:341-349(1999).
[8]
SEQUENCE OF 1-103 FROM N.A. (ISOFORM C), FUNCTION, AND TISSUE
SPECIFICITY.
RC TISSUE-Umbilical vein endothelial cells;
RX MEDLINE-98060745; PubMed-9396733;
RA Schweizer A., Valdenaire O., Nelboeck P., Deuschle U.,
Dumas Milne Edwards J.B., Stumpf J.G., Loeffler B.-M.,
RT "Human endothelin-converting enzyme (ECE-1): three isoforms with
distinct subcellular localizations.",
Biochem. J. 328:871-877(1997).
[9]
SEQUENCE OF 1-46 FROM N.A. (ISOFORM B).
RA Bagguely C.,
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
[10]
SEQUENCE OF 713-770 FROM N.A.
RA Flowers M.A., Tai S.C., Baluyut C.A., Cheung A.H., Kau C.L.,
Wong G.K.T., Marsden P.A.,
RT "Characterization of the human endothelin converting enzyme-1 gene
(ECE-1): genomic structure and chromosomal localization.",
Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CONVERTS BIG ENDOTHELIN-1 TO ENDOTHELIN-1.
CC -1- CATALYTIC ACTIVITY: Forms endothelin 1 by cleavage of the 21-
TTP-1-Val-22 bond in the precursor
CC -1- ENZYME REGULATION: INHIBITED BY PHOSPHORAMIDON.
CC -1- SUBUNIT: HOMODIMER.
CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; A, B (SHOWN HERE), C AND D; ARE
PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: ALL ISOFORMS ARE EXPRESSED IN UMBILICAL VEIN
ENDOTHELIAL CELLS, POLYNUCLEAR NEUTROPHILS, FIBROBLASTS, ATRIUM,
CARDIOMYOCYTES AND VENTRICLES. ISOFORMS A, B AND C ARE ALSO
EXPRESSED IN PLACENTA, LUNG, HEART, ADRENAL GLAND AND
PHAECHROMOCYTOMA; ISOFORMS A AND C IN LIVER, TESTIS AND SMALL
INTESTINE; ISOFORM B, C AND D IN ENDOTHELIAL CELLS AND UMBILICAL
VEIN SMOOTH MUSCLE CELLS; ISOFORMS C AND D IN SAPHENOUS VEIN
CELLS, AND ISOFORM C IN KIDNEY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M13 (ZINC
METALLOPROTEASE); ALSO KNOWN AS THE NEPRILYSIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
EMBL; 233307; CAA84548.1; ALT_INIT.
EMBL; D49471; BAA08442.1;
EMBL; X91922; CAA63015.1;
EMBL; X91923; CAA63015.1; JOINED.
EMBL; X91924; CAA63015.1; JOINED.
EMBL; X91925; CAA63015.1; JOINED.
EMBL; X91926; CAA63015.1; JOINED.
EMBL; X91927; CAA63015.1; JOINED.
EMBL; X91928; CAA63015.1; JOINED.
EMBL; X91929; CAA63015.1; JOINED.
EMBL; X91930; CAA63015.1; JOINED.
EMBL; X91931; CAA63015.1; JOINED.
EMBL; X91932; CAA63015.1; JOINED.
EMBL; X91933; CAA63015.1; JOINED.
EMBL; X91934; CAA63015.1; JOINED.
EMBL; X91935; CAA63015.1; JOINED.

EMBL; X91936; CAA63015.1; JOINED.
EMBL; X91937; CAA63015.1; JOINED.
EMBL; X91938; CAA63015.1; JOINED.
EMBL; X91939; CAA63015.1; JOINED.
EMBL; X91923; CAA63016.1;
EMBL; X91924; CAA63016.1; JOINED.
EMBL; X91925; CAA63016.1; JOINED.
EMBL; X91926; CAA63016.1; JOINED.
EMBL; X91927; CAA63016.1; JOINED.
EMBL; X91928; CAA63016.1; JOINED.
EMBL; X91929; CAA63016.1; JOINED.
EMBL; X91930; CAA63016.1; JOINED.
EMBL; X91931; CAA63016.1; JOINED.
EMBL; X91932; CAA63016.1; JOINED.
EMBL; X91933; CAA63016.1; JOINED.
EMBL; X91934; CAA63016.1; JOINED.
EMBL; X91935; CAA63016.1; JOINED.
EMBL; X91936; CAA63016.1; JOINED.
EMBL; X91937; CAA63016.1; JOINED.
EMBL; X91938; CAA63016.1; JOINED.
EMBL; X91939; CAA63016.1; JOINED.
EMBL; D43698; BAA07800.1;
EMBL; AB031742; BAA83687.1;
EMBL; AL031005; CAA19767.1;
EMBL; AJ130828; CAA64643.1;
EMBL; X98272; CAA66922.1;
EMBL; AL031728; CAA52285.1;
EMBL; AF018034; AAD21221.1;
HSP; P08473; IDMT.
MEROPS; M13.002;
MIM; 600423;
InterPro; IPR000718; Peptidase_M13.
InterPro; IPR000130; Zn_Mpeptidse.
Pfam; PF01431; Peptidase_M13; 1.
PRINTS; PR00786; NEPRILYSIN.
PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Glycoprotein; Transmembrane;
Signal-anchor; Alternative splicing; Polymorphism.
FT DOMAIN 1 68
CYTOPLASMIC (POTENTIAL)
SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL)
EXTRACELLULAR (POTENTIAL),
ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 607 608
BY SIMILARITY.
FT METAL 611 611
ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 667 667
PROTON DONOR (BY SIMILARITY).
FT CARBOHYD 671 671
N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 166 166
N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 187 187
N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 210 210
N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 270 270
N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 316 316
N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 362 362
N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 383 383
N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 539 539
N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 632 632
N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 651 651
N-LINKED (GLCNAC) (POTENTIAL).
FT VARSPLIC 1 44
MRGVPPPVPSALLSALGMSYKRAIDEDLDVDSISEGDY
PNG -> MPLOGIGLQRPFLQGRKPGGLTSPPPLLPSS
(IN ISOFORM A)
FT VARSPLIC 1 17
MRGVPPPVPSALLSALG -> MEALRESVHLALQ (IN
ISOFORM D).
FT VARSPLIC 1 17
MRGVPPPVPSALLSALG -> M (IN ISOFORM C).
FT VARIANT 341 341
T -> I (IN DBSNP:1076669).
/FTId=VAR_011972.
SQ SEQUENCE 770 AA; 87163 MW; DD88A59748B22F80 CRC64;

Query Match

55.3%; Score 367.5; DB 1; Length 770;

Best Local Similarity 53.4%; Pred. No. 1.4e-29;

Matches 62; Conservative 21; Mismatches 30; Indels 3; Gaps 1;

QY

1 GHEITHGFDDNGRNFKNMMDWWSNFSTQHFBSQSECMIYQYGNYSWDLADEQNNGF 60

-!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
-!- TISSUE SPECIFICITY: LYMPHOCTYCE AND FETAL BRAIN; NOT IN ADULT BRAIN, PLACENTA, SKELETAL MUSCLE, AND PANCREAS; NOT IN ADULT AND FETAL HEART, LUNG, LIVER, AND KIDNEY.
-!- DISEASE: DEFECTS IN PHEX ARE A CAUSE OF X-LINKED HYPOPHOSPHATEMIC RICKETS (HYP), A DOMINANT DISORDER CHARACTERIZED BY IMPAIRED PHOSPHATE UPTAKE IN THE KIDNEY, WHICH IS LIKELY TO BE CAUSED BY ABNORMAL REGULATION OF SODIUM PHOSPHATE COTRANSPORT IN THE PROXIMAL TUBULES. CLINICAL MANIFESTATIONS INCLUDE SKELETAL DEFORMITIES, GROWTH FAILURE, CRANIOSYNOSTOSIS, PARAVERTEBRAL CALCIFICATIONS, PSEUDOFRACTURES IN LOWER EXTREMITIES, AND MUSCULAR HYPOTONIA WITH ONSET IN EARLY CHILDHOOD. X-LINKED HYPOPHOSPHATEMIC RICKETS IS THE MOST COMMON FORM OF HYPOPHOSPHATEMIA WITH AN INCIDENCE OF 1 IN 20000.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M13 (ZINC METALLOPROTEASE); ALSO KNOWN AS THE NEPRILYSIN SUBFAMILY.
-!- DATABASE: NAME-PHEXdb. WWW="http://data.mch.mcgill.ca/phexdb/"

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

[illegible]

Query Match 53.1%; Score 353; DB 1; Length 749;
Best Local Similarity 55.2%; Pred. No. 3.9e-28;
Matches 64; Conservative 20; Mismatches 28; Indels 4; Gaps 2;

1 QY GHEITHGFDDNGRNFCKNGNMMDWWSNFTOHFREQSECMIVQYGNYSWDLADEQNNGE'60:
 ||| |||||:||| :||||| ||| | : |: :||| ||| ||| ||| |||
 db 579 GHEFTHGFDDNGRRKYDKNGNLDPWPSTESEKFKETKCMINOYSNYVKKAG-INVKKG 637

QY 61 NTLGENIADNGVROAYKAYLKWM---AEGGKDQQQLPGLDITHEQLFTFYNAVMC 113
| | | | | | | | | : | : | : | : | : | : | : |
Db 638 RTLGENIADNGLREAFRAYRKWINDRRQGLEPLPGTTFNNQLFSLVAHVRC 693

RESULT 10

PEX_MOUSE

ID PEX_MOUSE STANDARD; PRG: 749 AA.

AC P70669; P97439;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Metalloendopeptidase homolog PEX (EC 3.4.24.-) (Phosphate regulating neutral endopeptidase) (X-linked hypophosphatemia protein) (HYP)

DE Vitamin D-resistant hypophosphatemic rickets protein).

GZ PHOX OR PEX OR HYP.

OS Mus musculus (Mouse),

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;

[1]

RN SEQUENCE FROM N.A.

RX MEDLINE=96411643; PubMed=8812412;

RX Du L., Desbarats M., Viel J., Glorieux F.H., Cawthorn C., Ecarot B.; "cDNA cloning of the murine Pex gene implicated in X-linked hypophosphatemia and evidence for expression in bone."; Genomics 36:22-28(1996). [2]

RN SEQUENCE FROM N.A.

RX MEDLINE=97217775; PubMed=9063736;

RA Strom T.M., Francis F., Lorenz B., Boeddrich A., Econs M.J., Lehman H., Meitinger T.; "Pex gene deletions in Gy and Hyp mice provide mouse models for X-linked hypophosphatemia"; Hum. Mol. Genet. 6:165-171(1997). [3]

RN SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J;

RX MEDLINE=97232252; PubMed=9077527;

RA Beck L., Soumounou Y., Martel J., Krishnamurthy G., Gauthier C., Goodyer C.G., Tenenhouse H.S.; "Peyp/PEX tissue distribution and evidence for a deletion in the 3' region of the Pex gene in X-linked hypophosphatemic mice."; J. Clin. Invest. 99:1200-1209(1997).

CC -1- FUNCTION: PROBABLY INVOLVED IN BONE AND DENTIN MINERALIZATION AND RENAL PHOSPHATE REABSORPTION.

CC -1- SUBCELLULAR LOCATION: TYPE II membrane protein (Potential).

CC -1- TISSUE SPECIFICITY: BONE.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M13 (ZINC METALLOPROTEASE); ALSO KNOWN AS THE NEPRILYSIN SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licensee@isb-sib.ch).

EMBL; U49908; AAC36502.1; -
EMBL; U73910; AAC25962.1; -
EMBL; U73912; AAC25964.1; -
EMBL; U73913; AAC25965.1; -
EMBL; U73914; AAC25966.1; -
EMBL; U73911; AAC25963.1; -
EMBL; U73915; AAC25967.1; -
EMBL; U75646; ABA47750.1; -
HSSP: P08473; 1DMT
MEROPS; M13.091;
MGD; MGI:107489; Phex.
InterPro; IPR000718; Peptidase_M13.
InterPro; IPR000130; zn_MTpeptide.
Pfam; PF01431; Peptidase_M13; 1

DR PRINTS: PRO0786; NEPRILYSIN.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KW Hydrolase; Metalloprotease; Zinc; Glycoprotein; Transmembrane;
 FT DOMAIN 1 20
 FT TRANSMEM 21 37
 FT
 FT DOMAIN 38 749
 FT METAL 580
 FT ACT_SITE 581
 FT METAL 584
 FT METAL 584
 FT METAL 642
 FT ACT_SITE 646
 FT CARBOHYD 71
 FT CARBOHYD 238
 FT CARBOHYD 263
 FT CARBOHYD 290
 FT CARBOHYD 301
 FT CARBOHYD 377
 FT CARBOHYD 484
 FT CARBOHYD 736
 FT CONFLICT 430
 FT SEQUENCE 749 AA; A88FA481C376C18A CRC64;
 SQ
 Query Match 52.88; Score 351; DB 1; Length 749;
 Best Local Similarity 55.28; Pred. No. 6.1e-28;
 Matches 64; Conservative 20; Mismatches 28; Indels 4; Gaps 2;
 QY 1 GHEITHGFDNGRNFKNMMDWNSFSTQHFREQSECMYQYGNYSWDLADQNNGVF 60
 DB 579 GHEITHGFDNGRNFKNMMDWNSFSTQHFREQSECMYQYGNYSWDLADQNNGVF 60
 QY 61 NTLGENIADNGVROAYKAYLKWM---AEGKDDQPLGLDLTLPGLTFFINTYAVQWVC 113
 DB 638 RTLGENIADNGGLREAFRAYRKWINDRRQGVVEPLLPITTNQLPFLSAHVRC 693
 RESULT 11
 ID ECE2_HUMAN STANDARD; PRT; 765 AA.
 AC O60344;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 18-OCT-2001 (Rel. 40, Last annotation update)
 DE Endothelin-converting enzyme 2 (EC 3.4.24.71) (ECE-2).
 GN ECE2 OR KIAA0604.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RA "Prediction of the coding sequences of unidentified human genes. IX.
 RA The complete sequences of 100 new cDNA clones from brain which can
 RA code for large proteins in vitro.";
 RL DNA Res. 5:31-39(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lorenzo M.-N., Wang Y., Tai S.A., Chan G., Khan R.Y., Cheung A.H.,
 RA Marsden P.A.;
 RA "Human endothelin-converting enzyme 2 (ECE2): characterization of mRNA
 RA species and chromosomal localization.";
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CONVERTS BIG ENDOTHELIN-1 TO ENDOTHELIN-1 (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Forms endothelin 1 by cleavage of the 21-
 CC Trp-I-Val-22 bond in the precursor.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein.

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M13 (ZINC
 METALLOPROTEASE); ALSO KNOWN AS THE NEPRILYSIN SUBFAMILY.

 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; AB011176; BAA25530.1;
 DR EMBL; AF192531; AAG28399.1;
 DR HSP; P08473; 1DWT.
 DR MEROPS; M13.006;
 DR InterPro; IPR000718; Peptidase_M13.
 DR InterPro; IPR000130; Zn_MTPeptidse.
 DR Pfam; PF01431; Neprilase_M13; 1.
 DR PRINTS; PRO0786; NEPRILYSIN.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Hydrolase; Metalloprotease; Zinc; Glycoprotein; Transmembrane;
 KW Signal-anchor. 1 60
 KW DOMAIN 61 81
 FT TRANSMEM 61 81
 FT
 FT DOMAIN 82 765
 FT METAL 602
 FT ACT_SITE 603
 FT METAL 606
 FT METAL 662
 FT ACT_SITE 666
 FT CARBOHYD 161
 FT CARBOHYD 165
 FT CARBOHYD 206
 FT CARBOHYD 266
 FT CARBOHYD 311
 FT CARBOHYD 378
 FT CARBOHYD 534
 FT CARBOHYD 627
 FT CARBOHYD 635
 SQ SEQUENCE 765 AA; 86469 MW; 89D1B831B5628694 CRC64;
 Query Match 51.78; Score 343.5; DB 1; Length 765;
 Best Local Similarity 50.98; Pred. No. 3.6e-27;
 Matches 59; Conservative 18; Mismatches 36; Indels 3; Gaps 1;
 QY 1 GHEITHGFDNGRNFKNMMDWNSFSTQHFREQSECMYQYGNYSWDLADQNNGVF 60
 DB 601 GHEITHGFDNGRNFKNMMDWNSFSTQHFREQSECMYQYGNYSWDLADQNNGVF 60
 QY 61 NTLGENIADNGVROAYKAYLKWM---AEGKDDQPLGLDLTLPGLTFFINTYAVQWVC 116
 DB 658 QTLGENIADNGGLKAAAYKAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFPAQWCVSR 713
 RESULT 12
 ID ECE2_BOVIN STANDARD; PRT; 787 AA.
 AC O10711;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Endothelin-converting enzyme 2 (EC 3.4.24.71) (ECE-2).
 GN ECE2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95318093; PubMed=7797512;

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical zinc metalloproteinase T16A9.4 (EC 3.4.24.-).
GN T16A9.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA McMurray A.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RC STRAIN-BRISTOL N2;
RA Jones S.J.M.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (potential).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M13 (ZINC
CC METALLOPROTEASE); ALSO KNOWN AS THE NEPRILYSIN SUBFAMILY.
CC -----
CC CC

	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch .		
EMBL; Z77135;	CAB00879.1;	-	
HSP; P08473;	IDMT		
WormPep; T16A9.4;	CEI8259.		
InterPro; IPR000718;	Peptidase_M13.		
InterPro; IPR000130;	Zn_MTpeptdse.		
Pfam; PR01431;	Peptidase_M13; 1.		
PRINTS; PR00786;	NEPRILYSIN.		
PROSITE; PS00142;	ZINC_PROTEASE; 1.		
KW Hypothetical protein; Hydrolase;	Metalloprotease; Zinc; Glycoprotein;		
KW Transmembrane; Signal-anchor.			
DOMAIN	1	26	CYTOPLASMIC (POTENTIAL).
FT FT	27	47	SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).
FT DOMAIN	48	769	EXTRACELLULAR (POTENTIAL).
FT METAL	601.	601	ZINC (CATALYTIC (BY SIMILARITY)).
FT ACT_SITE	602	602	BY SIMILARITY.
FT METAL	605	605	ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL	663	663	ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE	667	667	PROTON DONOR (BY SIMILARITY).
FT CARBOHYD	69	69	N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD	221	221	N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD	240	240	N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD	272	272	N-LINKED (GLCNAC...) (POTENTIAL).

FT	CARBOHYD	356	356	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	412	412	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	506	506	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	684	684	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	698	698	N-LINKED (GLCNAC. . .)	(POTENTIAL)
SQ	SEQUENCE	769	AA; 88349	MW; EF6552CE49AA3928	CRC64;

Query Match 39.5%; Score 263; DB 1; Length 769;
 Best Local Similarity 44.0%; Pred. No. 4, 9e-19;
 Matches 51; Conservative 21; Mismatches 40; Indels 4; Gaps 3;

```

Db      600  GHEVSHAFDDGGGGYDGMGLNDMDWAETBEKTEKTRCFVRQYENVHVYVEAD-IHLNGQ  658
Qy      61  NTLGENTADNGVRQYKAYLKWMAE--GGKDQOLPGL-DLTHEQLQFFINQAVWC  113
Db      659  LSLGENTADNGGVKTAFNAYKAMKSNNTGTISEPALPGFQNFTSQMEFLAYANNWC  714

```

```
RESULT 14
PEPO_LACHE STANDARD; PRT; 647 AA.
AC 052071;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neutral endopeptidase (EC 3.4.24.-) (Endopeptidase O).
GN PEPO.
OS Lactobacillus helveticus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC Lactobacillus
OX NCBI_TaxID=1587;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CNR2 32;
RX MEDLINE=98395014; PubMed=9726890;
RA Chen Y.-S., Steele J.L.;
RT "Genetic characterization and physiological role of endopeptidase O
  from Lactobacillus helveticus CNR232."
RL Appl. Environ. Microbiol. 64:3411-3415(1998).
CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M13 (ZINC
  METALLOPROTEASE); ALSO KNOWN AS THE NEPRILYSIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL; AF019410; AAC35997.1;
DR MEROPS; M13.004;
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR000130; Zn_Mtpeptidse.
DR Pfam; PF01431; Peptidase_M13; 1.
DR PRINTS; PR00786; NEPRILYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc.
FT METAL 496 496 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 497 497 BY SIMILARITY.
FT METAL 500 500 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 556 556 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 560 560 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 647 AA; 73537 MW; EB5153C5E1DDE6FC CRC64;

Query Match 25.0%; Score 166.5; DB 1; Length 647;
Best Local Similarity 33.3%; Pred. No. 2.3e-09;
Matches 39; Conservative 19; Mismatches 32; Indels 27; Gaps 4;

QY 1 GHEITHGFDNGRFDKNGNMDWNSFSTQHFREQSECMY-----YQGNYSWDLADEQ 55
  |||::| |||::| |||::| |||::| |||::| |||::| |||::| |||::|
DB 495 GHEVSHAFDNGSGAKFDEHGNMNNWTDDEFAEFNKRQVMYDIFDGLQYG-----PA 546

QY 56 NVNGFNTLGENIADNGGVQRAYKAYLKWMAEGGKQQQLPGLDLTHEQLFFINYAQVW 112
  |||::| |||::| |||::| |||::| |||::| |||::| |||::| |||::|
DB 547 KINGQVGVGENIADLAGLACAVQ-----GRNDVLDKDL-----FENYARSW 589

RESULT 15
PEPO_LACHA STANDARD; PRT; 626 AA.
AC 007744;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neutral endopeptidase (EC 3.4.24.-) (Endopeptidase O).
GN PEPO OR L11803.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
```

```
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSL135;
RX MEDLINE=94064545; PubMed=8244921;
RA Tykkynen S., Buist G., Kunji E., Kok J., Poolman B., Venema G.,
  Haandrikman A.;
RT "Genetic and biochemical characterization of the oligopeptide
  transport system of Lactococcus lactis."
RL J. Bacteriol. 175:7523-7532(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ILI403;
RX MEDLINE=21235186; PubMed=1137471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
  Weissbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
  lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
CC -1- FUNCTION: ENDOPEPTIDASE WITH BROAD SUBSTRATE SPECIFICITY FOR
  SEVERAL OLIGOPEPTIDES. OPTIMUM ACTIVITY IS MEASURED AT
  PH 6.0-6.5 AND TEMPERATURES 30-38 DEGREES C.
CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M13 (ZINC
  METALLOPROTEASE); ALSO KNOWN AS THE NEPRILYSIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL; L18760; AAA16168.1;
DR MEROPS; M13.004;
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR000130; Zn_Mtpeptidse.
DR Pfam; PF01431; Peptidase_M13; 1.
DR PRINTS; PR00786; NEPRILYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Complete proteome.
FT INIT_MET 0 0 BY SIMILARITY.
FT METAL 474 474 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 475 475 BY SIMILARITY.
FT METAL 478 478 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 534 534 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 538 538 PROTON DONOR (BY SIMILARITY).
FT CONFLICT 75 75 T -> A (IN REF. 1).
FT CONFLICT 216 216 S -> N (IN REF. 1).
FT CONFLICT 350 350 A -> V (IN REF. 1).
FT CONFLICT 406 406 E -> K (IN REF. 1).
FT CONFLICT 470 470 T -> A (IN REF. 1).
SQ SEQUENCE 626 AA; 71406 MW; D89FB627D25F59FD CRC64;

Query Match 24.8%; Score 165; DB 1; Length 626;
Best Local Similarity 34.28; Pred. No. 3.1e-09;
Matches 38; Conservative 16; Mismatches 41; Indels 16; Gaps 3;

QY 2 HEITHGFDNGRFDKNGNMDWNSFSTQHFREQSECMYIYQGNYSWDLADEQNVNGFN 61
  |||::| |||::| |||::| |||::| |||::| |||::| |||::| |||::|
DB 474 HEISHAFDNGAQFDKGNLKNWLDDEYAEFEKQKEMIALFDGCVETEAGP---ANGKL 530

QY 62 TLGENIADNGGVQRAYKAYLKWMAEGGKQQQLPGLDLTHEQLFFINYAQVW 112
  |||::| |||::| |||::| |||::| |||::| |||::| |||::| |||::|
DB 531 IVSENIADQGGITTAALTA-----AKDEKVDL-----KAFFSQWAKIW 568
```

Mon Oct 7 11:24:40 2002

Search completed: October 3, 2002, 11:06:08
Job time: 41 sec

us-09-647-780a-4.rsp

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002; 11:05:27 ; Search time 12.99 seconds
(without alignments)
218.119 Million cell updates/sec.

Title: US-09-647-780A-4
Perfect score: 665
Sequence: 1 GHEITHGFDDNGRNFKNKN.....DLTHEQLFFINVAQVWCCK 116

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	374.5	56.3	189	3	US-08-646-273-14
2	374.5	56.3	708	3	US-08-646-273-23
3	374.5	56.3	754	3	US-08-646-273-30
4	370.5	55.7	758	1	US-08-289-112-2
5	367.5	55.3	703	3	US-08-646-273-25
6	367.5	55.3	753	3	US-08-646-273-36
7	339.5	51.1	787	1	US-08-574-763-2
8	336.5	50.6	567	3	US-08-646-273-19
9	318.5	47.9	775	4	US-05-305-640-2
10	73	11.0	758	1	US-08-258-188-2
11	73	11.0	758	1	US-08-526-813-2
12	73	11.0	758	5	PCT-US95-08554-2
13	70	10.5	999	2	US-08-473-553A-5
14	68	10.2	2710	2	US-08-568-459A-12
15	68	10.2	2710	2	US-08-487-826B-12
16	68	10.2	3060	2	US-08-487-826B-14
17	67	10.1	914	4	US-09-437-054A-8
18	65	9.8	500	3	US-08-926-842B-64
19	65	9.8	942	4	US-09-171-461-12
20	64.5	9.7	474	2	US-08-282-197C-58
21	64.5	9.7	2813	3	US-08-896-449A-2
22	64.5	9.7	2813	3	US-09-132-652-2
23	64	9.6	919	4	US-09-437-054A-17
24	63.5	9.5	2050	2	US-08-347-594A-2
25	63	9.5	214	3	US-08-960-507-21
26	63	9.5	478	4	US-08-740-223A-7
27	63	9.5	490	4	US-08-740-223A-12

28	63	9.5	495	4	US-08-740-223A-26
29	63	9.5	495	4	US-09-351-457-5
30	63	9.5	495	4	US-09-561-500-5
31	63	9.5	495	4	US-09-561-108-5
32	63	9.5	496	4	US-08-740-223A-22
33	63	9.5	497	1	US-08-373-579-4
34	63	9.5	497	2	US-08-418-595-4
35	63	9.5	497	2	US-08-665-926-4
36	63	9.5	497	2	US-08-348-492-4
37	63	9.5	497	4	US-09-162-437-4
38	63	9.5	497	4	US-08-740-223A-14
39	63	9.5	498	1	US-08-373-579-2
40	63	9.5	498	2	US-08-418-595-2
41	63	9.5	498	2	US-08-665-926-2
42	63	9.5	498	2	US-08-348-492-2
43	63	9.5	498	4	US-09-162-437-2
44	63	9.5	498	4	US-08-740-223A-2
45	63	9.5	498	4	US-09-351-457-2

ALIGNMENTS

RESULT 1
US-08-646-273-14
; Sequence 14, Application US/08646273
; Patent No. 6066502

GENERAL INFORMATION:
APPLICANT: Kroeger, Burkhard, Seulerberger, Harald, Meyer, Thomas, Schmidt, Heine
APPLICANT: Martin, Jacob, Elard, Otter, Rainer, Subkowski, Thomas, Hillen, Hein
TITLE OF INVENTION: Endothelin Converting Enzyme (ECE).
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kell & Weinlauf
STREET: 1101 Connecticut Avenue
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
COMPUTER: IBM AT-compatible, 80486 processor
OPERATING SYSTEM: MS-DOS version 6.0
SOFTWARE: WordPerfect version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,273
FILING DATE: 16-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/03706
FILING DATE: 11-NOV-1994
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 189 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-646-273-14

Query Match 56.3%; Score 374.5; DB 3; Length 189;

Best Local Similarity 54.3%; Pred. No. 5.4e-36;

Matches 63; Conservative 23; Mismatches 27; Indels 3; Gaps 1;

QY 1 GHEITHGFDDNGRNFKNKNMDFNSFQHFREQSECMYQYGNYSWDLADQNNGF 60

Db 25 GHEITHGFDDNGRNFKNKNMDFNSFQHFREQSECMYQYGNYSWDLADQNNGF 81

QY 61 NTGENIADNGVRQAYKAYLKWAEGKQDQPLGLDLTHEQLFFINVAQVWCCK 116

Db 82 HTLGENIADNGGLKAAAYRAYONVWVKNGAEQTLTLGLTNNQLFLSFAQVWCVR 137

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```


RESULT 9.
US-09-305-640-2
; Sequence 2, Application US/09305640B
; Patent No. 6255468
; GENERAL INFORMATION:

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-258-188-2

Query Match	11.08;	Score 73;	DB 1;	Length 758;
Best Local Similarity	24.28;	Pred. No. 4.1;		


```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,553A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Silva, Robin M.
REGISTRATION NUMBER: 38,304
REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 999 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-473-553A-5

Query Match 10.5%; Score 70; DB 2: Length 999;
Best Local Similarity 24.5%; Pred. No. 13;
Matches 26; Conservative 15; Mismatches 35; Indels 30; Gaps 6;

Qy 25 WSNFSTQHFREQ--SECMIYQVGNYSWDLADQNVNGFNTLGE--NIADNGGVROYAKAY 80
Db 666 WRSFKHLFSEHETADCL-----DEKNVIGFGSGKVKYKVELRGGEVAVKK- 712

Qy 81 LKMAEGGKDOQLFDLTHQLQLEFI-----NYAQWCCG 115
Db 713 LNKSVKGDDVEY--SSDSLNRDVFAAEVETLGTIRKHSIVRLWCCC 756

RESULT 14
US-08-568-459A-12
Sequence 12, Application US/08568459A
Patent No. 5849306
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellem, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/568,459A
FILING DATE: 07-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
```

```
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2710 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-08-568-459A-12

Query Match 10.2%; Score 68; DB 2: Length 2710;
Best Local Similarity 21.5%; Pred. No. 85;
Matches 23; Conservative 21; Mismatches 35; Indels 28; Gaps 5;

Qy 2 HEITHGFDONGNFDKNG-----NMMDWNSFSTQHFREQSECMIYQY----- 44
Db 1483 NDVAKADKIGKFFSKDGSKSPSLRSQEWKTKNGPEIKWGM-L-CALTRYVTDTDNKRKI 1541

Qy 45 -GNYSWDLADEQNVNGFNTLGENIADNGGVROYAKAYLKWMAEGGKD 90
Db 1542 KNDYSYDKVN-OSQNGNPSLEEPAK-----POFLRWMIWEGEE 1579

RESULT 15
US-08-487-826B-12
Sequence 12, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellem, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2710 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
```

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
US-08-487-826B-12

```

```

Query Match      10.2%; Score 68; DB 2; Length 2710;
Best Local Similarity 21.5%; Pred. No. 85;
Matches 23; Conservative 21; Mismatches 35; Indels 28; Gaps 5;

```

```

QY      2 HEITGFDNGRNFQKNG-----NMMDWSNFSTQHFREQSECMYQY----- 44

```

```

Db      1483 NDVAKAKDKIGKFFSKGSPGLSRQEWKTINGPEIWKGL-CALTRYVTDNKRKI 1541

```

```

QY      45 -GNYSWDLADEQNVNCFNTLGENIADNGGVRQAYKAYLKWMAEGGKD 90

```

```

Db      1542 KNDYSTYKVN-QSQNGNPSLEFAK-----POFLRMIEWGEE 1579

```

```

Search completed: October 3, 2002, 11:05:49
Job time: 22 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 11:05:27 ; Search time 24.7 Seconds.

(without alignments)

812.446 Million cell updates/sec

Title: US-09-647-780A-4

Perfect score: 665

Sequence: 1 GHEITHGFDDNGRNFKNKN.....DLTLEQLFFINVAQVWGCK 116

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 173994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

SPTREMBL_19.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	651	97.9	770	Q96PH9	Q96ph9 homo sapien
2	581	87.4	742	Q9ERK3	Q9erk3 mus musculus
3	581	87.4	765	Q9JLI3	Q9jli3 mus musculus
4	581	87.4	779	Q9ERK2	Q9erk2 mus musculus
5	577	86.8	742	Q9QZV6	Q9qzv6 mus musculus
6	577	86.8	765	Q9QZV7	Q9qzv7 mus musculus
7	434	65.3	770	Q93394	Q9q394 perca flave
8	397.5	59.8	661	Q9W435	Q9w435 drosophila
9	374.5	56.3	758	Q28868	Q28868 bos taurus
10	370.5	55.7	758	Q28010	Q28010 bos taurus
11	362.5	54.5	752	Q9DGN6	Q9dgn6 gallus gall
12	351	52.8	749	Q35812	Q35812 rattus norv
13	348.5	52.4	750	O44857	O44857 caenorhabdi
14	343.5	51.7	736	Q96NX4	Q96nx4 homo sapien
15	343.5	51.7	787	Q96NX4	Q96nx4 homo sapien
16	341.5	51.4	766	Q18673	Q18673 caenorhabdi

17	338.5	50.9	763	5	Q9XZ01-
18	338	50.8	848	5	O16796
19	336.5	50.6	763	11	Q923T6
20	323.5	48.6	193	5	Q95SM2
21	323	48.6	726	5	Q9W5Y0
22	321	48.3	772	5	Q9BLH1
23	318.5	47.9	775	4	Q9NY95
24	318.5	47.9	775	4	Q95672
25	313.5	47.1	770	5	Q9U9P2
26	306.5	46.1	775	11	Q9JHL3
27	306.5	46.1	775	11	Q9JMI0
28	300.5	45.2	787	5	Q9UA44
29	285	42.9	683	5	Q9Y136
30	285	42.9	700	5	Q9VAS0
31	274	41.2	717	5	Q9VAY0
32	257	38.6	754	5	O19831
33	253	38.0	1589	5	O45569
34	240.5	36.2	689	2	O50642
35	239.5	36.0	507	5	O9GTJ5
36	237	35.6	702	5	Q9VH96
37	230.5	34.7	837	5	O76751
38	228.5	34.4	682	5	Q9VAS2
39	223.5	33.6	700	16	Q9PFT1
40	222.5	33.5	793	5	Q25051
41	222.5	33.5	823	5	O45131
42	218.5	32.9	667	16	O06075
43	215.5	32.4	798	5	Q22763
44	208	31.3	870	5	Q9GTJ6
45	205	30.8	787	5	Q23684

ALIGNMENTS

RESULT 1
Q96PH9 PRELIMINARY: PRT; 770 AA.
ID Q96PH9
AC Q96PH9
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEPRILYSIN-LIKE METALLOPEPTIDASE 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21444797; PubMed=11560781;
RT Bonvouloir N., Lemieux N., Crine P., Boileau G., DesGroselliers L.;
RT "Molecular cloning, tissue distribution, and chromosomal localization
RT of WML2, a gene coding for a novel human member of the neutral
RT endopeptidase-24.11 family".
RL DNA Cell Biol. 20:493-498(2001).
DR EMBL; AF336981; AAL08942.1;
SQ SEQUENCE 770 AA; 88523 MW; DF9C08F4AA56A23B CRC64;

Query Match 97.9%; Score 651; DB 4; Length 770;
Best Local Similarity 100.0%; Pred. No. 4e-58;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY → 1 GHEITHGFDDNGRNFKNKNMMDWNSFSTQHFRESECMYQYGNYSNDLADBNVNGF 60
|||||
Db 603 GHEITHGFDDNGRNFKNKNMMDWNSFSTQHFRESECMYQYGNYSNDLADBNVNGF 662

QY 61 NTLGENIADNGVQYKAYLKWMAEGGKDDQLPGLDLTHEQLFFINVAQVWGCG 114
|||||
Db 663 NTLGENIADNGVQYKAYLKWMAEGGKDDQLPGLDLTHEQLFFINVAQVWGCG 716

RESULT 2
Q9ERK3

```
ID Q9ERK3 PRELIMINARY; PRT; 742 AA.
AC Q9ERK3; 2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE NEPRILYSIN-LIKE PEPTIDASE ALPHA.
GN MELL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21293028; PubMed=11278416;
RA Shirotani K., Tsubuki S., Iwata N., Takaki Y., Harigaya W.,
RA Maruyama K., Kiyu-Seo S., Kiyama H., Iwata H., Tomita T.,
RA Iwatsubo T., Saido T.C.;
RT "Neprilysin Degrades Both Amyloid beta Peptides 1-40 and 1-42 Most
RT Rapidly and Efficiently among Thiorphan- and Phosphoramidon-sensitive
RT Endopeptidases."
RL J. Biol. Chem. 276:21895-21901(2001).
DR EMBL; AF302075; AAG18446.1;
DR HSSP; P08473; 1DMT.
DR MEROPS; M13.008;
DR MGD; MGI:1351603; Mell1.
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR000130; Zn_Mtpeptdse.
DR Pfam; PF01431; Peptidase_M13; 1.
DR PRINTS; PR00786; NEPRILYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 742 AA; 85993 MW; 4A44EAD211B2499F CRC64;

Query Match 87.4%; Score 581; DB 11; Length 742;
Best Local Similarity 85.1%; Pred. No. 5.5e-51;
Matches 97; Conservative 14; Mismatches 3; Indels 0; Gaps 0;

QY 1 GHEITHGDDNGRNFCKNGNMMDWNSFSTQHFQSECMYQYGNYSWDLADQNNGVF 60
DB 575 GHEITHGDDNGRNFCKNGNMMDWNSFSTQHFQSECMYQYGNYSWDLADQNNGVF 634
QY 61 NTLGENIADNGGVQRAYKAYLKWMAEGGKQDQLPGLDLTHEQLFFINYAQVWCG 114
DB 635 STLGENIADNGGVQRAYKAYLWLDGKQDQLPGLNLTYAQVWCG 688

RESULT 3
Q9JLI3
ID Q9JLI3 PRELIMINARY; PRT; 765 AA.
AC Q9JLI3; 2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE NEPRILYSIN-LIKE METALLOPEPTIDASE 1 (NEPRILYSIN-LIKE PEPTIDASE
DE BETA).
GN MELL1 OR SEP OR NL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=TESTIS;
RX PubMed=10749671;
RX Ghaddar G., Ruchon A.F., Carpentier M., Marcinkiewicz M., Seidah N.G.,
RX Crine P., DesGroselliers L., Boileau G.;
RT "Molecular cloning and biochemical characterization of a new mouse
RT testis soluble zinc-metalloproteinase of the neprilysin family."
RL Biochem. J. 347:419-429(2000).
DR EMBL; AF302075; AAG18446.1;
DR HSSP; P08473; 1DMT.
DR MEROPS; M13.008;
DR MGD; MGI:1351603; Mell1.
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR000130; Zn_Mtpeptdse.
DR Pfam; PF01431; Peptidase_M13; 1.
DR PRINTS; PR00786; NEPRILYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 765 AA; 85993 MW; 4A44EAD211B2499F CRC64;

Query Match 87.4%; Score 581; DB 11; Length 742;
Best Local Similarity 85.1%; Pred. No. 5.5e-51;
Matches 97; Conservative 14; Mismatches 3; Indels 0; Gaps 0;

QY 1 GHEITHGDDNGRNFCKNGNMMDWNSFSTQHFQSECMYQYGNYSWDLADQNNGVF 60
DB 575 GHEITHGDDNGRNFCKNGNMMDWNSFSTQHFQSECMYQYGNYSWDLADQNNGVF 634
QY 61 NTLGENIADNGGVQRAYKAYLKWMAEGGKQDQLPGLDLTHEQLFFINYAQVWCG 114
DB 635 STLGENIADNGGVQRAYKAYLWLDGKQDQLPGLNLTYAQVWCG 688

RESULT 3
Q9JLI3
ID Q9JLI3 PRELIMINARY; PRT; 765 AA.
AC Q9JLI3; 2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE NEPRILYSIN-LIKE METALLOPEPTIDASE 1 (NEPRILYSIN-LIKE PEPTIDASE
DE BETA).
GN MELL1 OR SEP OR NL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=TESTIS;
RX PubMed=10749671;
RX Ghaddar G., Ruchon A.F., Carpentier M., Marcinkiewicz M., Seidah N.G.,
RX Crine P., DesGroselliers L., Boileau G.;
RT "Molecular cloning and biochemical characterization of a new mouse
RT testis soluble zinc-metalloproteinase of the neprilysin family."
RL Biochem. J. 347:419-429(2000).
DR EMBL; AF302075; AAG18446.1;
DR HSSP; P08473; 1DMT.
DR MEROPS; M13.008;
DR MGD; MGI:1351603; Mell1.
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR000130; Zn_Mtpeptdse.
DR Pfam; PF01431; Peptidase_M13; 1.
DR PRINTS; PR00786; NEPRILYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 765 AA; 85993 MW; 4A44EAD211B2499F CRC64;

Query Match 87.4%; Score 581; DB 11; Length 779;
Best Local Similarity 85.1%; Pred. No. 5.8e-51;
Matches 97; Conservative 14; Mismatches 3; Indels 0; Gaps 0;

QY 1 GHEITHGDDNGRNFCKNGNMMDWNSFSTQHFQSECMYQYGNYSWDLADQNNGVF 60
DB 612 GHEITHGDDNGRNFCKNGNMMDWNSFSTQHFQSECMYQYGNYSWDLADQNNGVF 671
QY 61 NTLGENIADNGGVQRAYKAYLKWMAEGGKQDQLPGLDLTHEQLFFINYAQVWCG 114
```

```
Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF176569; AAG69247.1;
DR EMBL; AF302076; AAG18447.1;
DR HSSP; P08473; 1DMT.
DR MEROPS; M13.008;
DR MGD; MGI:1351603; Mell1.
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR000130; Zn_Mtpeptdse.
DR Pfam; PF01431; Peptidase_M13; 1.
DR PRINTS; PR00786; NEPRILYSIN.
DR PROSITE; PS00294; PRENYLIATION; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 765 AA; 88699 MW; D3662F1CE5B957F7 CRC64;

Query Match 87.4%; Score 581; DB 11; Length 765;
Best Local Similarity 85.1%; Pred. No. 5.7e-51;
Matches 97; Conservative 14; Mismatches 3; Indels 0; Gaps 0;

QY 1 GHEITHGDDNGRNFCKNGNMMDWNSFSTQHFQSECMYQYGNYSWDLADQNNGVF 60
DB 598 GHEITHGDDNGRNFCKNGNMMDWNSFSTQHFQSECMYQYGNYSWDLADQNNGVF 657
QY 61 NTLGENIADNGGVQRAYKAYLKWMAEGGKQDQLPGLDLTHEQLFFINYAQVWCG 114
DB 658 STLGENIADNGGVQRAYKAYLWLDGKQDQLPGLNLTYAQVWCG 711

RESULT 4
Q9ERK2
ID Q9ERK2 PRELIMINARY; PRT; 779 AA.
AC Q9ERK2; 2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE NEPRILYSIN-LIKE PEPTIDASE GAMMA.
GN MELL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21293028; PubMed=11278416;
RX Shirotani K., Tsubuki S., Iwata N., Takaki Y., Harigaya W.,
RX Maruyama K., Kiyu-Seo S., Kiyama H., Iwata H., Tomita T.,
RX Iwatsubo T., Saido T.C.;
RT "Neprilysin Degrades Both Amyloid beta Peptides 1-40 and 1-42 Most
RT Rapidly and Efficiently among Thiorphan- and Phosphoramidon-sensitive
RT Endopeptidases."
RL J. Biol. Chem. 276:21895-21901(2001).
DR EMBL; AF302077; AAG18448.1;
DR HSSP; P08473; 1DMT.
DR MEROPS; M13.008;
DR MGD; MGI:1351603; Mell1.
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR000130; Zn_Mtpeptdse.
DR Pfam; PF01431; Peptidase_M13; 1.
DR PRINTS; PR00786; NEPRILYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 779 AA; 89709 MW; 5E48CA530828584B CRC64;

Query Match 87.4%; Score 581; DB 11; Length 779;
Best Local Similarity 85.1%; Pred. No. 5.8e-51;
Matches 97; Conservative 14; Mismatches 3; Indels 0; Gaps 0;

QY 1 GHEITHGDDNGRNFCKNGNMMDWNSFSTQHFQSECMYQYGNYSWDLADQNNGVF 60
DB 612 GHEITHGDDNGRNFCKNGNMMDWNSFSTQHFQSECMYQYGNYSWDLADQNNGVF 671
QY 61 NTLGENIADNGGVQRAYKAYLKWMAEGGKQDQLPGLDLTHEQLFFINYAQVWCG 114
```

```

Db 672 STILGNIADNGGVQRAYKAYLRLWLDGGKQDRLPGLNLTVAQLFFINYAQVWCG 725
RESULT 5
Q90ZV6 PRELIMINARY; PRT; 742 AA.
AC Q90ZV6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SOLUBLE SECRETED ENDOPEPTIDASE DELTA.
GN MELL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20011457; PubMed=10542292;
RA Ikeda K., Emoto N., Raharjo S.B., Nurhantari Y., Saiki K.,
RA Yokoyama M., Matsuo M.;
RT "Molecular identification and characterization of novel membrane-bound
RT metalloprotease, the soluble secreted form of which hydrolyzes a
RT variety of vasoactive peptides."
RL J. Biol. Chem. 274:32469-32477(1999).
DR EMBL; AF157106; AAF13153.1;
DR HSSP; P08473; 1DWT.
DR MEROPS; M13.008;
DR MGD; MGI:1351603; Meil1.
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF01431; Peptidase_M13; 1.
DR PRINTS; PR00786; NEPRILYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 742 AA; 85945 MW; CF5FF9D982471157 CRC64;

Query Match 86.8%; Score 577; DB 11; Length 742;
Best Local Similarity 84.2%; Pred. No. 1.4e-50;
Matches 96; Conservative 15; Mismatches 3; Indels 0; Gaps 0;

QY 1 GHEITHGFDNGRNFKNMGMDWNSFNSTQHFREQSECMIYQYGNYSWDLADQNNGVF 60
Db 575 GHEITHGFDNGRNFKNMGMDWNSFNSTQHFREQSECMIYQYGNYSWDLADQNNGVF 634
QY 61 NTLGNIADNGGVQRAYKAYLRLWLDGGKQDRLPGLNLTVAQLFFINYAQVWCG 114.
Db 635 SSLGNIADNGGVQRAYKAYLRLWLDGGKQDRLPGLNLTVAQLFFINYAQVWCG 688

RESULT 6
Q90ZV7 PRELIMINARY; PRT; 765 AA.
AC Q90ZV7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SOLUBLE SECRETED ENDOPEPTIDASE.
GN MELL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20011457; PubMed=10542292;
RA Ikeda K., Emoto N., Raharjo S.B., Nurhantari Y., Saiki K.,
RA Yokoyama M., Matsuo M.;
RT "Molecular identification and characterization of novel membrane-bound
RT metalloprotease, the soluble secreted form of which hydrolyzes a
RT variety of vasoactive peptides."
RL J. Biol. Chem. 274:32469-32477(1999).

```

```

DR EMBL; AF157105; AAF13152.1;
DR HSSP; P08473; 1DWT.
DR MEROPS; M13.008;
DR MGD; MGI:1351603; Meil1.
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF01431; Peptidase_M13; 1.
DR PRINTS; PR00786; NEPRILYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 765 AA; 88651 MW; 567D3C17764C0F3F CRC64;

Query Match 86.8%; Score 577; DB 11; Length 765;
Best Local Similarity 84.2%; Pred. No. 1.5e-50;
Matches 96; Conservative 15; Mismatches 3; Indels 0; Gaps 0;

QY 1 GHEITHGFDNGRNFKNMGMDWNSFNSTQHFREQSECMIYQYGNYSWDLADQNNGVF 60
Db 598 GHEITHGFDNGRNFKNMGMDWNSFNSTQHFREQSECMIYQYGNYSWDLADQNNGVF 657
QY 61 NTLGNIADNGGVQRAYKAYLRLWLDGGKQDRLPGLNLTVAQLFFINYAQVWCG 114
Db 658 SSLGNIADNGGVQRAYKAYLRLWLDGGKQDRLPGLNLTVAQLFFINYAQVWCG 711

RESULT 7
Q93394 PRELIMINARY; PRT; 770 AA.
ID Q93394;
AC Q93394;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NEPRILYSIN.
GN NEP1.
OS Perca flavescens (Yellow perch).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Percidae; Perca.
OX NCBI_TaxID=8167;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-OVULATORY OVARY.
RX MEDLINE=99445407; PubMed=10514552;
RA Langenau D.M., Goetz F.W., Roberts S.B.;
RT "The upregulation of messenger ribonucleic acids during 17alpha,
RT 20beta-dihydroxy-4-pregnen-3-one-induced ovulation in the perch
RT ovary."
RL J. Mol. Endocrinol. 23:137-152(1999).
DR EMBL; AF077612; AAC28366.1;
DR HSSP; P08473; 1DWT.
DR MEROPS; M13.001;
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF01431; Peptidase_M13; 1.
DR PRINTS; PR00786; NEPRILYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 770 AA; 88135 MW; D29F216CB041BACC CRC64;

Query Match 65.3%; Score 434; DB 13; Length 770;
Best Local Similarity 62.3%; Pred. No. 6.1e-36;
Matches 71; Conservative 25; Mismatches 18; Indels 0; Gaps 0;

QY 1 GHEITHGFDNGRNFKNMGMDWNSFNSTQHFREQSECMIYQYGNYSWDLADQNNGVF 60
Db 603 GHEITHGFDNGRNFKNMGMDWNSFNSTQHFREQSECMIYQYGNYSWDLADQNNGVF 662
QY 61 NTLGNIADNGGVQRAYKAYLRLWLDGGKQDRLPGLNLTVAQLFFINYAQVWCG 114
Db 663 NTLGNIADNGGVQRAYKAYLRLWLDGGKQDRLPGLNLTVAQLFFINYAQVWCG 716

```

```
RESULT 8
Q9W435 PRELIMINARY; PRT: 661 AA.
AC Q9W435;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CG5894 PROTEIN.
GN CG5894.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris K.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laso P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003436; AAF46125.1;
DR HSSP; P08473; 1DWT.
DR MEOPS; M13 UPW;
DR FlyBase; FBgn0029843; Nepl.
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR000130; Zn_Mtpeptidse.
DR Pfam; PF01431; Peptidase_M13; 1.
DR PRINTS; PR00786; NEPRILysin.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
SQ SEQUENCE 661 AA; 76226 MW; 76DD13FC3BC066A5 CRC64;
```

```
Query Match 59.8%; Score 397.5; DB 5; Length 661;
Best Local Similarity 59.6%; Pred. No. 2.8e-32;
Matches 68; Conservative 21; Mismatches 24; Indels 1; Gaps 1;
```

```
QY 1 GHEITHGFDGNGRFDKNGNMDWNSFSTQHFRESECMYGYNYSWDLADQNNGVF 60
```

```
Db 495 GHEITHGFDGNGRFDKNGNMDWNSFSTQHFRESECMYGYNYSWDLADQNNGVF 553
QY 61 NTLGENIADNGGVROYAKYKWKMAEGGKDDQLPGLDLTHEQLFFINIAQVCG 114
Db 554 MTOGENIADNGGLKQAFRAYKKWETLHGREQQLPGLNMTHTDQLFLFYAQIWC 607

RESULT 9
Q28868 PRELIMINARY; PRT: 758 AA.
AC Q28868;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ENDOTHELIN CONVERTING ENZYME.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RX MEDLINE=95032010; PubMed=7945289;
RA Ikura T., Sawamura T., Shiraki T., Hosokawa H., Kido T., Hoshikawa H.,
RA Shinada K., Tanzawa K., Kobayashi S., Miwa S., et al.;
RT "cDNA cloning and expression of bovine endothelin converting enzyme.";
RL Biochem. Biophys. Res. Commun. 203:1417-1422(1994).
DR EMBL; S73774; AAB32062.1;
DR HSSP; P08473; 1DWT.
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR000130; Zn_Mtpeptidse.
DR Pfam; PF01431; Peptidase_M13; 1.
DR PRINTS; PR00786; NEPRILysin.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
SQ SEQUENCE 758 AA; 85648 MW; 6A7EA96566BEFEC5 CRC64;

Query Match 56.3%; Score 374.5; DB 6; Length 758;
Best Local Similarity 54.3%; Pred. No. 7.3e-30;
Matches 63; Conservative 23; Mismatches 27; Indels 3; Gaps 1;

QY 1 GHEITHGFDGNGRFDKNGNMDWNSFSTQHFRESECMYGYNYSWDLADQNNGVF 60
Db 594 GHEITHGFDGNGRFDKNGNMDWNSFSTQHFRESECMYGYNYSWDLADQNNGVF 650
QY 61 NTLGENIADNGGVROYAKYKWKMAEGGKDDQLPGLDLTHEQLFFINIAQVCG 116
Db 651 HTLGENIADNGGLKAAAYRAYQYQYVKNVKNKGAETLPTLTGNTNQLFSLFAQVWC 706

RESULT 10
Q28010 PRELIMINARY; PRT: 758 AA.
AC Q28010;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ENDOTHELIN CONVERTING ENZYME-1A.
GN ECE-1A.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RX MEDLINE=94340737; PubMed=8062389;
RA Xu D., Emoto N., Giald A., Slaughter C., Kaw S., deWit D.,
RA Yanagisawa M.;
RT "ECE-1: a membrane-bound metalloprotease that catalyzes the
RT proteolytic activation of big endothelin-1";
RL Cell 78:473-485(1994).
```

```

RN  [2]
RP  SEQUENCE FROM N.A.
RA  Yanagisawa M.;
RL  Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
DR  EMBL; U27342; AAA82928.1; -
DR  HSP; P08473; IDMT.
DR  MEROPS; M13.002; -
DR  InterPro; IPR000718; Peptidase_M13.
DR  InterPro; IPR000130; Zn_MTPeptdse.
DR  Pfam; PF01431; Peptidase_M13; 1.
DR  PRINTS; PR00786; NEPRILYSIN.
DR  PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
SQ  SEQUENCE 758 AA; 85620 MW; 39DCB702099F40A4 CRC64;

Query Match
Best Local Similarity 55.7%; Score 370.5; DB 6; Length 758;
Matches 62; Conservative 23; Mismatches 28; Indels 3; Gaps 1;

QY  1 GHEITHGFDNGRNFDRKNGNMMDWNSFSTQHFREQSECMYQYGNYSWDLADQNVNGF 60
Db  594 GHEITHAFDDQGRYDKDGNLRPMWKNSSVEAFKQTACMVQYGNYS---VNGEPVNGR 650
QY  61 NTLGENTADNGVRQAYKAYLKWMAGGKQDQQLPGLDLTHEOLFNFYNAQVWCCK 116
Db  651 HTLGENIADNGGLKAARYAYQVWVKNGAEOTLTLGLTNHOLFNFYNAQVWCCK 706

RESULT 11
Q9DGN6
ID  Q9DGN6 PRELIMINARY; PRT; 752 AA.
AC  Q9DGN6;
DT  01-MAR-2001 (TREMBLrel. 16, Created)
DT  01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT  01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE  ENDOTHELIN CONVERTING ENZYME-1.
GN  ECE-1.
OS  Gallus gallus (Chicken).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC  Gallus.
OX  NCBI_TaxID=9031;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  MEDLINE-20363660; PubMed-10903177;
RA  Takebayashi-Suzuki K., Yanagisawa M., Gourdle R.G., Kanzawa N.,
RA  Mikawa T.;
RT  "In vivo induction of cardiac Purkinje fiber differentiation by
RT  coexpression of preproendothelin-1 and endothelin converting enzyme-
RT  1.";
RL  Development 127:3523-3532(2000).
DR  EMBL; AF230274; AAF98287.1; -
DR  HSP; P08473; IDMT.
DR  MEROPS; M13.002; -
DR  InterPro; IPR000718; Peptidase_M13.
DR  InterPro; IPR000130; Zn_MTPeptdse.
DR  Pfam; PF01431; Peptidase_M13; 1.
DR  PRINTS; PR00786; NEPRILYSIN.
DR  PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
SQ  SEQUENCE 752 AA; 84985 MW; 3B6B3DB7BF900656 CRC64;

Query Match
Best Local Similarity 54.5%; Score 362.5; DB 13; Length 752;
Matches 60; Conservative 24; Mismatches 29; Indels 3; Gaps 1;

QY  1 GHEITHGFDNGRNFDRKNGNMMDWNSFSTQHFREQSECMYQYGNYSWDLADQNVNGF 60
Db  588 GHEITHAFDDQGRYDKDGNLRPMWKNSSVEAFKQTACMVQYGNYS---VNGEPVNGR 650
QY  61 NTLGENTADNGVRQAYKAYLKWMAGGKQDQQLPGLDLTHEOLFNFYNAQVWCCK 116
Db  645 HTLGENIADNGGLKAARYAYQVWVKNGAEOTLTLGLTNHOLFNFYNAQVWCCK 700
```

```

RESULT 12
Q35812
ID  Q35812 PRELIMINARY; PRT; 749 AA.
AC  Q35812;
DT  01-JAN-1998 (TREMBLrel. 05, Created)
DT  01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT  01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE  PHEX PROTEIN.
GN  PHEX.
OS  Rattus norvegicus (Rat).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX  NCBI_TaxID=10116;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  TISSUE-INCISOR;
RA  MEDLINE-97343325; PubMed-9199930;
RA  Francis F., Strom T.M., Hennig S., Boeddrich A., Lorenz B.,
RA  Brandau O., Mohrke K.L., Cagnoli M., Steffens C., Klages S.,
RA  Borzym K., Pohl T., Oudet C., Econs M.J., Rowe P.S., Reinhardt R.,
RA  Weisinger T., Lehrach H.;
RT  "Genomic organization of the human PEX gene mutated in X-linked
RT  dominant hypophosphatemic rickets.";
RL  Genome Res. 7:573-585(1997).
DR  EMBL; AJ001637; CAA04890.1; -
DR  HSP; P08473; IDMT.
DR  MEROPS; M13.091; -
DR  InterPro; IPR000718; Peptidase_M13.
DR  InterPro; IPR000130; Zn_MTPeptdse.
DR  Pfam; PF01431; Peptidase_M13; 1.
DR  PRINTS; PR00786; NEPRILYSIN.
DR  PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
SQ  SEQUENCE 749 AA; 86305 MW; 690B8E2609C1A63 CRC64;

Query Match
Best Local Similarity 52.8%; Score 351; DB 11; Length 749;
Matches 64; Conservative 20; Mismatches 28; Indels 4; Gaps 2;

QY  1 GHEITHGFDNGRNFDRKNGNMMDWNSFSTQHFREQSECMYQYGNYSWDLADQNVNGF 60
Db  579 GHEITHGFDNGRKYDKNGNDLPWVSDEKFKETKCMINYSNTYWKAG-LNVKRG 637
QY  61 NTLGENTADNGVRQAYKAYLKWM---AEGKQDQQLPGLDLTHEOLFNFYNAQVWC 113
Db  638 RTLGENIADNGGLREAFRAYRWKINDRQGVPEPLPGITFTNNQLFLSTAHRVC 693

RESULT 13
O44857
ID  O44857 PRELIMINARY; PRT; 750 AA.
AC  O44857;
DT  01-JUN-1998 (TREMBLrel. 06, Created)
DT  01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT  01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE  HYPOTHETICAL 87.0 KDA PROTEIN.
GN  T0528.4.
OS  Caenorhabditis elegans.
OC  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC  Rhabditidae; Peloderinae; Caenorhabditis.
OX  NCBI_TaxID=6239;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  STRAIN-BRISTOL N2;
RA  MEDLINE-99069613; PubMed-9851916;
RA  None;
RT  "Genome sequence of the nematode C. elegans: a platform for
RT  investigating biology. The C. elegans Sequencing Consortium.";
RL  Science 282:2012-2018(1998).
RN  [2]
RP  SEQUENCE FROM N.A.
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 3, 2002, 11:05:27 ; Search time 16.05 Seconds
(without alignments)
694,477 Million cell updates/sec.

Title: US-09-647-780A-4
Perfect score: 665
Sequence: 1 GHEITHGDDNGRNFKNKGN.....DLTHEQLFFINYAQVWCGCK 116

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	578	86.9	774	JC7265	neprilysin (EC 3.4.24.11)
2	469	70.5	750	HYHUN	neprilysin (EC 3.4.24.11)
3	489	70.5	751	HYRIN	neprilysin (EC 3.4.24.11)
4	485	69.9	750	HYRIN	neprilysin (EC 3.4.24.11)
5	374.5	56.3	754	S47268	endothelin convert
6	370.5	55.7	754	A53679	endothelin convert
7	370.5	55.7	758	A54667	endothelin convert
8	367.5	55.3	758	JC2521	endothelin convert
9	367.5	55.3	770	JC4136	endothelin convert
10	348.5	52.4	750	D88082	protein T05A8.4 [i]
11	341.5	51.4	766	T20003	hypothetical prote
12	339.5	51.1	825	I46078	endothelin convert
13	338	50.8	890	C88099	protein Fl8A12.8 [i]
14	263	39.5	769	T24949	hypothetical prote
15	257	38.6	754	T16182	hypothetical prote
16	253	38.0	1589	T22668	hypothetical prote
17	223.5	33.6	700	B82788	hypothetical prote
18	218.5	32.9	667	C87236	metalloproteinase X
19	215.5	32.4	798	T28906	probable zinc meta
20	205	30.8	823	T28132	hypothetical prote
21	202.5	30.5	706	F87683	hypothetical prote
22	201.5	30.3	663	T20838	peptidase M13 fami
23	179	26.9	630	F95191	probable zinc meta
24	179	26.9	630	C88058	endopeptidase O [i]
25	165	24.8	627	C86850	endopeptidase O (E
26	165	24.8	627	A47098	neutral endopeptid
27	163	24.8	627	F53290	endopeptidase Pepo
28	163.5	24.6	564	T11548	probable zinc meta
29	163	24.5	534	T32020	hypothetical prote

protein Fl8A12.1 [i]
hypothetical prote
protein Fl8A12.4 [i]
protein Fl8A12.3 [i]
kell blood group p
protein Fl8A12.6 [i]
hypothetical prote
hypothetical prote
hyaluronate lyase
hypothetical prote
chitinase homolog
large repetitive p
uncharacterized pr
zinc metalloprotei
hypothetical prote
hypothetical prote

ALIGNMENTS

RESULT 1

JC7265

neprilysin (EC 3.4.24.11) II - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 02-Sep-2000

C:Accession: JC7265

R:Janja, O.; Facchinetti, P.; Rose, C.; Bonhomme, M.C.; Gros, C.; Schwartz, J.C.

Biochem. Biophys. Res. Commun. 271, 565-570, 2000

A:Title: Neprilysin II: A putative novel metalloprotease and its isoforms in CNS and

A:Reference number: JC7265

A:Contents: Brain and testis

A:Accession: JC7265

A:Molecule type: mRNA

A:Residues: 1-774 <TAN>

C:Genetics:

A:Gene: nepII

C:Superfamily: neprilysin

C:Keywords: brain; glycoprotein; hydrolase; metalloproteinase; neuropeptide; testis;

RESULT 2

HYHUN

neprilysin (EC 3.4.24.11) [validated] - human

N:Alternate names: Cdi0; common acute lymphocytic leukemia antigen; endopeptidase 24.

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 08-Dec-2000

C:Accession: A41387; A36173; S05275; J10084; S00350; S02228

P:O'Donoghue, J.; Shropshire, J.; Masteller, E.L.; Clayton, E.L.

Proc. Natl. Acad. Sci. USA 89:4819-4823, 1992

A:Title: Molecular cloning of the common acute lymphoblastic leukemia antigen (CALLA)

A:Reference number: A41387; MUID:89386688

A:Accession: A41387

A:Molecule type: DNA

A:Residues: 1-750 <DAA>

A:Cross-references: GB:M26605

A:Note: the authors translated the codon AAC for residues 14 and 72 as Asp.

E.L. Clayton, E.L. Masteller, E.L. Clayton, E.L. Clayton

Proc. Natl. Acad. Sci. USA 89:4819-4823, 1992

A:Title: Molecular cloning of the common acute lymphoblastic leukemia antigen (CALLA)

A:Reference number: A41387; MUID:89386688

A:Accession: A41387

A:Molecule type: DNA

A:Residues: 1-750 <DAA>

A:Cross-references: GB:M26605

A:Note: the authors translated the codon AAC for residues 14 and 72 as Asp.

E.L. Clayton, E.L. Masteller, E.L. Clayton, E.L. Clayton

Proc. Natl. Acad. Sci. USA 89:4819-4823, 1992

A:Title: Molecular cloning of the common acute lymphoblastic leukemia antigen (CALLA)

A:Reference number: A41387; MUID:89386688

A:Accession: A41387

A:Molecule type: DNA

A:Residues: 1-750 <DAA>

A:Cross-references: GB:M26605

A:Note: the authors translated the codon AAC for residues 14 and 72 as Asp.

E.L. Clayton, E.L. Masteller, E.L. Clayton, E.L. Clayton

Proc. Natl. Acad. Sci. USA 89:4819-4823, 1992

A:Title: Molecular cloning of the common acute lymphoblastic leukemia antigen (CALLA)

A:Reference number: A41387; MUID:89386688


```
F/52-750/Domain: extracellular #status predicted <EXT>
F/145,285,325,628/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/311/Binding site: carbohydrate (Asn) (covalent) #status experimental
F/584,588/Binding site: zinc (His) #status predicted
F/585/Active site: Glu #status predicted.
```

Query Match 69.9%; Score 465; DB 1; Length 750;
Best Local Similarity 64.9%; Pred. No. 5.7e-37;
Matches 74; Conservative 26; Mismatches 14; Indels 0; Gaps 0;

Qy	1	GHEITHGFDNDGRNFKNGNMMDWNSFSTQHFRSEQSECMYQYGNYSWDLADEQNNGVF	60
		: : :: : : : : : : : :	
Db	583	GHEITHGFDNDGRNFNKDGLVDVWTQQSANFNKDKSQCMYYQYGNFTWDLAGGOHLNGI	642

Qy	61	NTLGENIADNGGVRQAYKAYLKWMAEGGQQPLGLDITHEQLFFINYAQVWC	114
		: :	
		: :	
Db	643	NTLGENIADNGGIGQAYRAYQYVKKNGEKLLPGLDLNHKOLFFLNFAOVWC	696

RESULT 5
S47268
endothelin-converting enzyme (EC 3.4.24.-) 1 - bovine
C/Species: Bos primigenius taurus (cattle)
C/Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text change 08-Oct-1999

R.; Schmidt, M.; Kroeger, B.; Jacob, E.; Seulberger, H.; Subkowski, T.; Otter, R.; Meyer, F.E.B. Lett. 356, 238-243, 1994
A; title: Molecular characterization of human and bovine endothelin converting enzyme (EC 3.4.21.57)
A; Reference number: S51009; MUID:95104423

A: Cross-references: EMBL:Z35306; NID:G935074; PID:CAA84547.1; PID:G535075
C: Superfamily: neprilysin
C: Keywords: hydrolase; metalloproteinase; zinc
F: 591.595/Binding site: zinc, catalytic (His) #status predicted
F: 592/Active site: Glu #status predicted

Query Match 56.3%; Score 374.5; DB 2; Length 754;
Best Local Similarity 54.3%; Pred. NO. 3e-28;
Matches 63; Conservative 23; Mismatches 27; Indels 3; Gaps 1;

QY 1 GHEITHGDDNGRNFDKGNMMDWNSFNSTOHFREQSECMYGYGNSVDLADQNUNGF 60
||||| |||| | :||: || | : :||: ||: ||||| : |||
Db 590 GHETHAFEDDGRKYDKGNLRPPWNKSSVEAFKOOTACMVEOYGNS---VNGRPVNGR 646

Qy 61 NTLGENIADNGVROAYKAYALKWMAEGGKDQQLPLGLDTHEQLFFINYAQWCGCK 116
:|||||: |||| : : | | | | : |||| :
Dd 647 HTLGENIADNGLKAARYAYONVYKKNGAEOTLP TLGTNNOLFELSAFQWCVSVR 702

RESULT 6
A53679
endothelin converting enzyme (EC 3.4.24.-) - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-Oct-1994 #sequence revision 07-Oct-1994
#text change 20-Jun-2000

R; Shimada, K.; Takahashi, M.; Tanzawa, K.
J. Biol. Chem. 269, 18275-18278, 1994
A; Title: Cloning and functional expression of endothelin-converting enzyme from rat
A; Reference number: A53679; MUID: 94308046

A; Residues: 1-734 SHS>
A; Cross-references: GB:D29683; NID:g529084; PIDN:BAA06152.1; PID:g529085
C: Superfamily: neprilysin
C: Keywords: glycoprotein; hydrolase; metalloproteinase; transmembrane protein
F: 1-52/Domain: intracellular #status predicted <N>
F: 53-72/Domain: extracellular #status predicted <N>

F;53- / 3/Domain: transmembrane #status: predicted <TM>

```
F:74-754/Domain: extracellular #status predicted <EXT>
F:129/Binding site: substrate (Arg) #status predicted
F:150,171,194,254,300,346,367,523,616,635/Binding site: carbohydrate (Asn) (covalent)
F:591,595/Binding site: zinc, catalytic (His) #status predicted
F:592/Active site: Glu #status predicted
```

Query Match	55.7%;	Score 370.5;	DB 2;	Length 754;
Best Local Similarity	54.3%;	Pred. NO. 7.4e-28;		
Matches 63:	Conservative	21;	Mismatches 29;	Indels 3;
				Gaps 1;

QY 1 GHEITHGFDNCRNFKNGNMDDWSNFSTQHFEQSECHYYQYNSWDLADEQN VNGF 60
|||:|||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

Dd 590 GHEITHAFDQGGREYDKDGNIURPWWKSSVEAFKOOTECWVOOYNYS---VNGEPVNGR 616

QY 61 NTLGENIADNGVROAYKAYLKWMAEGGKDOQLPGDLTHEQLFFINYAQWCGCK 116
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
db 647 HTLGENIADNGLKAAVRAYONVYWKNGAEOTLPTTGITSNOLFGEAOWCVR 702

```

RESULT 7
A54667
endothelin converting enzyme (EC 3.4.24.-) 1 - bovine
C:Species: Bos primigenius taurus (Cattle)
C:Date: 28-Oct-1994 sequence_revision 28-Oct-1994 #text_change 05-Nov-1999
C:Accession: A54667; JC2448
C:Xu, D.; Emoto, N.; Glaid, A.; Slaughter, C.; Kaw, S.; dewit, D.; Yanagisawa, M.
Cell 78, 473-485, 1994
A:Title: ECE-1: a membrane-bound metalloprotease that catalyzes the proteolytic cleavage of endothelin-1
A:Reference number: A54667; GUID:94340737

```

A; Residues: 1-736 <A>
A; Cross-references: GB:U27342; NID:9897601; PIDN:AA83928.1; PID:9897602.
R; Kura, T.; Sawamura, T.; Shiraki, T.; Hosokawa, H.; Kido, T.; Hoshikawa, H.; Shima
Biochem. Biophys. Res. Commun. 203, 1417-1422, 1994
A; Title: cDNA cloning and expression of bovine endothelin converting enzyme.
A; Reference number: JG2448. MID:95032010

A: Molecule type: mRNA
A: Residues: 1-90, 'A', '92-698', 'A', '700-758' <TKU>
A: Cross-references: GB|S57377; NID:9688289; PIDN:AAB32062.1; PID:9688290
C: Superfamily: neprilysin
C: Keywords: hydrolase; metalloproteinase; transmembrane protein; zinc
F: 5-77/Domain: transmembrane #status predicted <TKM>
F: 595-599/Binding site: zinc, catalytic (His) #status predicted
F: 596/Active site: Glu #status predicted

Query Match	Score	DB 2	Length
Best Local Similarity	55.7%	370.5	758
Matches	53.4%	Pred. No. 7.4e-28	
Conservative	23	Mismatches	28
Conservative	23	Indels	33
Gaps	1		

QY 1 GHEITHGFDNGRNDKNGNMDDWNSFSTQHFREQSECMYQYGNYSWDLADEQNNGVF 60
|||:||| ||| :|||: ||| :|||: ||| :|||: ||| :|||
db 594 GHEITHAFDODGREYDKDGNIRPWWNSSVEAFKOOTACMVEQVGNYS---VNFGNCR 650

QY 61 NTLGENIADNGVQRAYKAYLKWMAEGKKDQQLPGLDTHEQLFFINYAQVWCCK 116
 :|||||:||||: |||| | : : ||| ||| :|||||: |||| :
Db 651 HTLGENIADNGLKAARYQNWKKGAEQTPTLGLTNNQLNFSLFSGQVWCYSVR 706

RESULT 8
JC2521
endothelin converting enzyme (EC 3.4.24.-).1, umbilical vein endothelial cell form
C:Species: Homo sapiens (man)
C:date: 17-May-1995 #sequence_revision 14-Jul-1995 #text_change 20-Jun-2000
C:Accession: JC2521

R;Shimada, K.; Matsushita, Y.; Wakabayashi, K.; Takahashi, M.; Matsubara, A.; Iijima, Biochem. Biophys. Res. Commun. 207, 807-812, 1995
A;Title: Cloning and functional expression of human endothelin-converting enzyme cDNA
A;Reference number: JCS251; MUID:95169128

A:Accession: JC2521
A:Molecule type: mRNA
A:Residues: 1-758 <SHI>
A:Cross-references: GB:D43698; NID:g1197803; PIDN:BAA07800.1; PID:g1197804
A:Experimental source: umbilical vein endothelial cells
C:Genetics:
A:Gene: GDB:ECE1; ECE
A:Cross-references: GDB:698357; OMIM:600423
A:Map position: lp36.1-lp36.1
C:Superfamily: neprilysin
C:Keywords: alternative splicing; glycoprotein; hydrolase; metalloproteinase; transmembrane protein
F:57-77/Domain: transmembrane #status predicted <TM>
F:154,175,198,258,304,350,371,527,620,639/Binding site: carbohydrate (Asn) (covalent)
F:595,599/Binding site: zinc, catalytic (His) #status predicted
F:596/Active site: Glu #status predicted

Query Match 55.3%; Score 367.5; DB 2; Length 758;
Best Local Similarity 53.4%; Pred. No. 1.4e-27;
Matches 62; Conservative 21; Mismatches 30; Indels 3; Gaps 1;

OY 1 CHEITHGFDNGRNFEDKGNMMDWNSFTQHFREQSECMIYQXGYNSWDLADENVNGF 60
||||| ||||| ||||| : ||||| : ||||| : ||||| : |||||
Db 594 GHETHAFDDGGREYDKDGLNRPMKKSSVEAFARQEPCMVESYNY--VVGPEYNGR 650
: ||||| ||||| ||||| : ||||| : ||||| : ||||| : |||||
OY 61 NTLGENTADNGGVROAYKAYLKMAEGGKDQLPGLDLTHEQLFFINVAOWWCCK 116
:||||| ||||| ||||| : ||||| : ||||| : ||||| : |||||
Db 651 HTLGENTADNGGLKAARYAVONYWVKKGAEHSLEPTLGITNNOLFELFGFAQCWSVR 706
: ||||| ||||| ||||| : ||||| : ||||| : ||||| : |||||

RESULT 9.
JC4136
endothelin converting enzyme (EC 3.4.24.-) 1, renal adenocarcinoma form - human
C:Species: Homo sapiens (man)
C>Date: 31-Aug-1995 #sequence_revision 27-Oct-1995 #text_change 20-Jun-2000
C:Accession: JC4136; S51009; S47269
R:Yorimitsu, K.; Moroi, K.; Inagaki, N.; Saito, T.; Masuda, Y.; Masaki, T.; Selno, S.; M
Biochem. Biophys. Res. Commun. 208, 721-727, 1995
A>Title: Cloning and sequencing of a human endothelin converting enzyme in renal adeno
A:Reference number: JC4136; MUID:95209687
A:Accession: JC4136
A:Molecule type: mRNA
A:Residues: 1-770 <YOR>
A:Cross-references: GB:D49471; NID:g695404; PIDN:BAA08442.1; PID:g695405
A:Experimental source: renal adenocarcinoma cell
A>Note: it is uncertain whether Met-1 or Met-18 is the initiator
R:Schmidt, M.; Kroeger, B.; Jacob, E.; Seuberger, H.; Subkowski, T.; Otter, R.; Meyer,
FEBS Lett. 356, 238-243, 1994
A>Title: Molecular characterization of human and bovine endothelin converting enzyme (EC
A:Reference number: S51009; MUID:95104423
A:Accession: S51009
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 18-770 <SC2>
A:Cross-references: EMBL:D235307; NID:g535181; PIDN:CAA84548.1; PID:g535182
C:Comment: This enzyme is a phosphoramidon-sensitive, membrane-bound metalloprotease.
C:Genetics:
A:Gene: GDB:ECE1; ECE
A:Cross-references: GDB:698357; OMIM:600423
A:Map position: lp36.1-lp36.1
C:Superfamily: neprilysin
C:Keywords: alternative splicing; glycoprotein; hydrolase; metalloproteinase; phosphop
F:63-89/Domain: transmembrane #status predicted <TM>
F:25/Binding site: phosphate (Thr) (covalent) #status predicted
F:166,187,210,270,316,323,383,539,632,651/Binding site: carbohydrate (Asn) (covalent)
F:607,611/Binding site: zinc, catalytic (His) #status predicted
F:608/Active site: Glu #status predicted

Query Match 55.3%; Score 367.5; DB 2; Length 770;
Best Local Similarity 53.4%; Pred. No. 1.5e-27;
Matches 62; Conservative 21; Mismatches 30; Indels 3; Gaps 1;

